Tissue-specific ChIP-seq of Drosophila Salivary Gland Transcription Factors Reveal that CrebA Directly Regulates Secretory Regulators and Effectors

by

Dorothy M. Johnson

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
August, 2018

ABSTRACT

Transcription factors (TFs) regulate the expression of tissue specific genes. In the Drosophila embryonic salivary gland (SG), four of these transcription factors include: *CrebA*, *fork head* (*fkh*), *Sage*, and *sensless* (*sens*). CrebA and its Creb3-like homologs are known for their regulation of secretory capacity. *fkh* is a regulator of SG morphogenesis and, with *Sage* and *sens*, controls SG survival. With the exception of *Sage*, all of these TFs are expressed in multiple tissues. To determine the SG-specific direct targets of these TFs, all four were tagged with GFP in preparation for Chromatin Immunoprecipitation and sequencing (ChIP-seq) in which GFP would be pulled down. The GFP-tagged forms were expressed using two drivers that have the SG in common, so by studying the intersection of these two datasets, SG-specific binding can be inferred. As controls, wild-type embryonic chromatin is immunoprecipitated with either CrebA or Sage antisera. For CrebA, which is ubiquitously expressed, this potentially permits us to distinguish tissue-specific targets of CrebA from CrebA targets bound in all tissues. Because Sage is SG specific, a comparison of the GFP datasets and the wild-type Sage pull down dataset controls for the accuracy ofgood the Sage-GFP dataset. All of the GFP tagged TFs are fully functional, and to date, we have obtained the results for CrebA binding. For the Fkh, Sage, and Sens data, we have sent the prepared chromatin to our collaborators at University of Minnesota, Duluth. The CrebA ChIP-seq datasets revealed that CrebA binds to open chromatin containing consensus CrebA binding sites; that the subset of genes that CrebA regulates are involved in secretion, including *Xbp1*, *TSN*, and *tbc1*; and that loss of CrebA and/or the binding sites leads to misregulation of *Xbp1* and *TSN*. The highly conserved gene, *tbc1* is known to be involved in the trans Golgi network from mammalian studies and we find that it is localized to the trans most portion of the TGN in Drosophila. Loss of *tbc1* leads to apical membrane defects and secretion defects in the SG and cuticle secretion defects. This predicted Rab-GAP co-localizes with several Rabs but does not affect the steady-state levels two of these Rabs, Rab5 and Rab11.

ACKNOWLEDGEMENTS

This PhD, like all, has been a journey. Thank you to all who have helped me prepare and finish my mine.

To my husband, Fr. Chris: From when I started at Hopkins we have had three children, two ordinations (all yours), and one fire. You've been with me during my highs and the lows, always with a word of encouragement. I'm unbelievably grateful to have you by my side (and to be by yours).

To my children, Daniel, Charlotte, and Justinian: You make sure I get a break from work (whether I like it or not) and there is nothing in this world so awesome as coming home to a chorus of "Mama!"s, or, in the case of Justinian, smiles.

To my parents: Thank you for being there for me from my first breath in this world through now. There is nothing like knowing you have support regardless of what you do. To my mom, especially, thank you for watching the kids. You have been the stop-gap on days when no one else could watch them and I needed to go in to lab. Without your help, I wouldn't have been able to finish. To my dad: you started all this the day we went to the train tracks and brought home tadpoles to watch their metamorphosis into frogs. You have always encouraged us to learn as much as possible about any subject, even if that meant taking your eldest daughter to the Smithsonian to look at a bunch of rocks (well, minerals).

To my siblings Andrew, Grace, Peter, Catherine, and Tim, and my two sister-in-laws Bethanie and Naomi: I can't imagine life without you guys. It is always fun to hang out and be able to relax around you guys. Even if Andrew wins at [insert any game here].

To my in-laws, Fr. George and Mat. Deborah: Thank you for welcoming me into the family with open arms. I have never doubted for a moment that you see me as your daughter. Also, thank you for watching the kids. I don't know how we would have been able to make it to this point if you didn't.

To my mentor Debbie: From the moment I joined the lab, I knew it was the place for me. Your love of science and excitement to see new results seems as fresh as the day you started. You have always emphasized to "go with what the data says" and, even when the data seems to be indicating the opposite of what we'd thought, you can always think of it in a positive light.

To the past and present Andrew Lab members: Raj, Mike, JiHoon, Cait, Parama, Afshan, Yim, Rika, Emily, Kyla, Se Yeon, Rebecca, Aria and others: thank you for the helpful comments in lab meeting, for answering any questions, and helping us win multiple departmental contests.

Thank you to my thesis committee: Dr. Susan Michaelis, Dr. Brendan Cormack, Dr. Joel Pomerantz and Dr. Caroline Machamer.

To SangJoon Kim and Abigail Bastien: You both helped me in multiple ways. I have not forgotten SangJoon's cheerful, sing-song greeting in the morning. Without your comments, I would not have discovered that UAS-fkh-GFP is male infertile. Abigail: You were instrumental in performing the experiments in Figure 8 and Table 1. You both are missed greatly. May your memory be eternal! +

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION:

CHAPTER 1:

CHAPTER 2:

PAGE INTENTIONALLY TO BE BLANK

INTRODUCTION

Tubular organs are responsible for many vital processes, including the exchange of gases (vertebrate lungs, insect trachea), absorption and secretion (kidneys, small intestine, pancreas, salivary glands), and circulation of nutrients and immune cells (blood vessels) (Andrew and Ewald, 2010). Some organs, such as the central nervous system, initially start development as a tube, but their further development obscures this early architecture. Many diseases are directly related to issues in tube formation and/or maintenance, including spina bifida, polycystic kidney disease, and aortic aneurysms. Thus, understanding the cellular and molecular mechanisms underlying the formation, homeostasis and function of tubular organs is vital. Direct study of mammalian tubular organs is complicated by the intricacy of the organs, redundant gene function, and *in utero* development. On the other hand, the salivary gland (SG) of *Drosophila melanogaster* is an ideal model system for the study of epithelial tube biology because it forms a simple, unbranched tubular epithelium in a matter of only a few hours. Embryos are clear and immuno-stained SGs are easily visible by light and confocal microscopy. An armamentarium of genetic tools exists in Drosophila for eliminating gene function, either ubiquitously or in a tissue-specific manner, or for overexpressing various forms of different proteins in specific cell types (Kanca et al., 2017). Finally, genetic redundancy is much less of a factor in flies than in mammals (Yamamoto et al., 2014; Chung et al., 2014).

Salivary gland specification is evident early in development (embryonic stage 10) as two SG placodes (plates of columnar epithelial cells) on the ventral surface of the embryo in the posterior head region begin to express SG-specific proteins (Chung et al., 2014). Each placode contains around 140 cells and, after specification, there is no cell division or cell death as this tissue develops. Once specified, SG cells change shape, rearrange and invaginate dorsally into the embryo to form simple, unbranched epithelial tubes (embryonic stages 11 and 12) (Chung et al., 2017; Roper, 2012). As the cells of the developing SG tubes contact the layer of muscle surrounding the developing gut, known as the visceral mesoderm, the tube turns and migrates

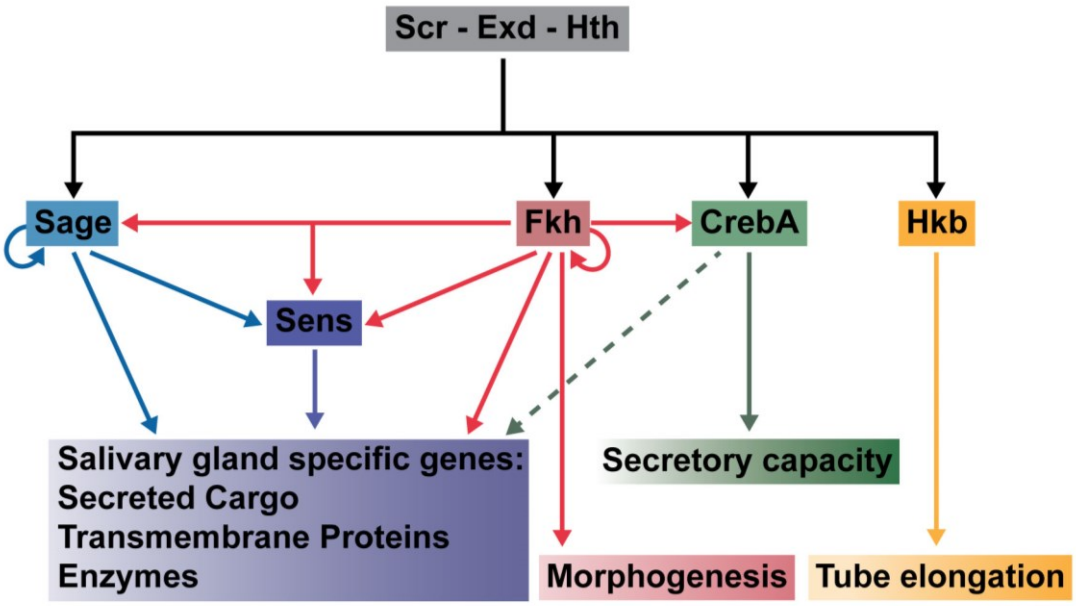


Figure 1: TF Pathway of Drosophila Embryonic Salivary Gland. Modified from Fox et al 2013

posteriorly along this structure to its final position, parallel to the long axis of the embryo (stages 13 - 15) (Bradley and Andrew, 2001). SGs invaginate and migrate as fully polarized, intact epithelial tubes, although the SG cells undergo considerable rearrangement to transform the SG primordia into internalized, elongated tubular structures (Sanchez-Corrales et al., 2018).

The embryonic SG placode is specified by the combined activities of three transcription factors (TFs): *Sex combs reduced* (*Scr*), *Extradenticle* (*Exd*), and *Homothorax* (*Hth*) (Figure 1) (Henderson and Andrew, 2000). These TFs function as a complex to activate the expression of several other TFs involved in the survival, morphogenesis, and function of the SG: *fork head* (*fkh*), *sage*, *CrebA*, and *huckebein* (*hkb*). Fkh and Sage, in turn, activate the expression of the TF *senseless* (*sens*) (Fox et al., 2013). Shortly after expression of Fkh, Sage, CrebA, and Hkb are activated and as SG morphogenesis begins, expression of Scr, Exd, and Hth disappears. Fkh, Sage and CrebA continue to be expressed in the SG throughout the lifetime of the organ, with Fkh maintaining its own expression and expression of CrebA, Sage and Sens. Hkb, on the other hand, is only transiently expressed in the early SG.

Loss of *fkh*, *sage*, or *sens* leads to loss of the embryonic SG through apoptotic cell death (Chandrasekaran and Beckendorf, 2003; Myat and Andrew, 2000a). SG apoptosis can be rescued by the presence of a deficiency that removes the apoptotic regulators that are activated in the SG when these proteins are missing (Myat and Andrew, 2000a; Fox et al., 2013) or overexpression of the baculovirus P35 anti-apoptotic protein (Fox et al., 2013). Sens expression in the SG requires both *fkh* and *sage* and the ectopic activation of *sens* requires both Fkh and Sage (Fox et al., 2013). All of the genes encoding SG specific cargo and the enzymes that modify that cargo require Fkh and Sage, and often Sens, for full levels of SG expression, and the ectopic expression of Fkh and Sage is sufficient to activate expression of these SG-specific genes in ectopic sites in the embryos (Fox et al., 2013). Stainings of 3rd instar larval SG polytene chromosomes with antisera to Fkh, Sage and Sens reveal significant overlap in the binding sites for all three proteins. Binding sites for these three TFs do not, for the most part, overlap with the binding sites observed for CrebA (Maruyama et al., 2011;Fox et al., 2013). This led to the hypothesis that Fkh, Sage, and Sens act in a complex to regulate both SG survival and expression of SG specific gene products. CrebA is proposed to function in parallel to increase secretory capacity (Fox et al., 2010).

When the cell death of *fkh* mutant SGs is rescued, the SGs have severe morphological defects: the SGs completely fail to undergo the cell shape changes associated with early SG development and remain as two placodes of cells on the embryo surface (Weigel et al., 1989). Rescuing the cell death phenotype of *Sage* mutants results in relatively normal SGs, with a significant reduction in lumenal volume, consistent with the decreases in SG cargo gene expression that are also observed.

Unlike the other four TFs, Hkb, a Sp1/Egr-like TF, is only transiently expressed in the SG (Myat and Andrew, 2002). SG expression of *hkb* mRNA normally ceases during invagination. Loss of *hkb* leads to a failure in apical expansion due to defects in targeting of apical vesicles, resulting in SGs that are fully internalized but not elongated (Myat and Andrew, 2000b). Not

surprisingly, persistent overexpression of Hkb leads to overly expanded apical lumens with fully elongated SG tubes (Myat and Andrew, 2002).

CrebA is unique among the early expressed TFs because its loss results in only mild morphological SG defects (Abrams and Andrew, 2005). Rather than playing a role in SG morphogenesis, CrebA is critical for the function of the SG. Specifically, CrebA increases secretory capacity in the SG by coordinate upregulation of expression of Secretory Pathway Component Genes (SPCGs) (Abrams and Andrew, 2005; Fox et al., 2010). This was initially shown by *in situs* comparing expression of 34 known SPCGs in wild-type and *CrebA* mutant embryos and by subsequent microarray analyses comparing mRNA levels in wild-type and *CrebA* mutant embryos. The microarray analysis revealed that CrebA not only upregulates SPCGs but also expression of mRNAs encoding SG cargo. It is unclear if the upregulation of SG cargo by CrebA is direct or indirect but it is likely that CrebA could directly activate most SPCGs. Fox et al (2010) showed that SG expression of several (4/5) tested SPCGs required the presence of consensus CrebA binding sites (Fox et al., 2010). A universal role for the CrebA family in the regulation of SPCGs is supported by experiments showing that expression of CrebA or the activated forms of any of the five human homologs of CrebA (known as the Creb3-like proteins) is sufficient to induce ectopic expression of all tested SPCGs in Drosophila embryos (Fox et al., 2010; Barbosa et al., 2013). In further support of CrebA/Creb3L proteins functioning to boost secretory capacity is the observation that murine loss of the two most closely related CrebA homologs, Creb3L1/OASIS and Creb3L2/BBF7H2, have osteogenesis and chondrogenesis defects, respectively, and both proteins have been implicated in upregulating expression of a subset of COPII proteins (Keller et al, 2017; Saito et al 2009).

Sage is the only known TF expressed solely in the SG (Abrams et al., 2006). Fkh and CrebA are expressed in multiple tissues beside the SG, including in the proventriculus, hindgut, epidermis and central nervous system (Andrew et al., 1997; Weigel et al., 1989). Sens is also expressed in the peripheral nervous system, whereas *hkb* is also expressed in a subset of cells in

the central nervous system (Fox et al., 2013; Myat and Andrew, 2000b). This widespread expression of SG TFs presents a challenge for assaying tissue-specific binding of these TFs by ChIP-Seq. To get around the issue of the broad expression domains of most of these proteins, we tagged each of the TFs with GFP to perform ChIP-seq, pulling down on SG expressed TF-GFP. CrebA, Fkh, Sage, and Sens tagged with GFP were expressed in the SG using two different drivers. Neither driver is SG specific, but the only tissue in which both are expressed is the SG. Thus, we can identify SG specific binding events by determining which peaks overlap in independent experiments using the two drivers. The ChIP-seq and analysis of CrebA is complete (discussed in Chapter 1), whereas we are awaiting the ChIP-seq data from our collaborators for Fkh, Sage, and Sens (discussed in Chapter 3).

In Chapter 1, we find that CrebA is binding open chromatin in the SG and that binding is driven by the presence of known consensus CrebA binding motifs determined by *in vitro* assays (Abrams and Andrew, 2005; Nitta et al., 2015; Noyes et al., 2008). We show that CrebA binds almost all SPCGs but only a small subset of SG cargo genes. In addition to SPCGs, CrebA also activates expression of other regulators of gene expression, Including *Xbp1* and *TSN*, a transcription factor and potential regulator of mRNA stability, respectively. CrebA also regulates a novel SPCG, *tbc1*, which has been shown to be involved in endosome to Golgi trafficking in mammalian systems (Shin et al., 2017). Mutation of CrebA binding sites in *Xbp1* or *TSN*, led to cuticle secretion defects. Loss of *TSN* or *tbc1* caused similar phenotypes. In the SG, mutations in all three genes resulted in secretion defects. Altogether, this suggests that CrebA not only regulates effectors of secretory capacity (such as *tbc1*) but also regulators (such as *Xbp1* and *TSN*).

Further analysis of *tbc1*, discussed in Chapter 2, shows that loss of *tbc1* leads to apical membrane irregularities. GFP-tagged Tbc1 does not localize to secretory vesicles or the ER, but does partially colocalize with the recycling endosome marker Rab11 near the apical surface. Endogenous Tbc1 also partially colocalizes with the *trans*-Golgi marker Golgin-245 but not the

cis-Golgi marker GM130. Tbc1 is predicted to be a Rab-GAP (GTPase Activating Protein). To determine any potential interactors, Tbc1 and candidate YFP-Rabs were overexpressed mosaically in the SG. Tbc1 partially colocalized with Rab-X1, -5, -8, -9, and -10. In an effort to determine if Tbc1 might have an effect on the levels of Rab proteins, we tested Rab5 and Rab11 and found that neither loss nor overexpression of Tbc1 had an overt effect on Rab5 or Rab11 accumulation levels.

In Chapter 3, the generation and testing of the GFP constructs for Fkh, Sage, and Sens are discussed. It was found that ubiquitous expression of each tagged protein phenocopies the activities observed with ubiquitous expression of the corresponding untagged forms. Expression of the GFP tagged TFs also rescued the SG cell death phenotypes associated with loss of each of the TFs. Finally, the Sage antiserum that will be used as a control (pulling down Sage in wild type embryos) fully recognizes Sage-GFP. Altogether, these experiments suggest that the GFP constructs are fully functional. Chromatin for these experiments has been prepared and shipped to our collaborators at the University of Minnesota, Duluth. We excitedly await the results from this analysis.

References

Abrams, E.W., and Andrew, D.J. (2005). CrebA regulates secretory activity in the Drosophila salivary gland and epidermis. Development *132,* 2743-2758.

Abrams, E.W., Mihoulides, W.K., and Andrew, D.J. (2006). Fork head and Sage maintain a uniform and patent salivary gland lumen through regulation of two downstream target genes, PH4aSG1 and PH4aSG2. Development *133,* 3517-3527.

Andrew, D.J., Baig, A., Bhanot, P., Smolik, S.M., and Henderson, K.D. (1997). The Drosophila dCREB-A gene is required for dorsal/ventral patterning of the larval cuticle. Development *124,* 181-193.

Andrew, D.J., and Ewald, A.J. (2010). Morphogenesis of epithelial tubes: Insights into tube formation, elongation, and elaboration. Dev. Biol. *341,* 34-55.

Barbosa, S., Fasanella, G., Carreira, S., Llarena, M., Fox, R., Barreca, C., Andrew, D., and O'Hare, P. (2013). An orchestrated program regulating secretory pathway genes and cargos by the transmembrane transcription factor CREB-H. Traffic *14,* 382-398.

Bradley, P.L., and Andrew, D.J. (2001). *ribbon* encodes a novel BTB/POZ protein required for directed cell migration in *Drosophila melanogaster*. Development *128,* 3001-3015.

Chandrasekaran, V., and Beckendorf, S.K. (2003). senseless is necessary for the survival of embryonic salivary glands in Drosophila. Development *130,* 4719-4728.

Chung, S., Hanlon, C.D., and Andrew, D.J. (2014). Building and specializing epithelial tubular organs: the Drosophila salivary gland as a model system for revealing how epithelial organs are specified, form and specialize. Wiley Interdisciplinary Reviews: Developmental Biology *3,* 281-300.

Chung, S., Kim, S., and Andrew, D.J. (2017). Uncoupling apical constriction from tissue invagination. ELife *6,* e22235.

Fox, R.M., Hanlon, C.D., and Andrew, D.J. (2010). The CrebA/Creb3-like transcription factors are major and direct regulators of secretory capacity. The Journal of Cell Biology *191,* 479-492.

Fox, R.M., Vaishnavi, A., Maruyama, R., and Andrew, D.J. (2013). Organ-specific gene expression: the bHLH protein Sage provides tissue specificity to Drosophila FoxA. Development *140,* 2160-2171.

Henderson, K.D., and Andrew, D.J. (2000). Regulation and Function of Scr, exd, and hth in the Drosophila Salivary Gland. Dev. Biol. *217,* 362-374.

Kanca, O., Bellen, H.J., and Schnorrer, F. (2017). Gene Tagging Strategies To Assess Protein Expression, Localization, and Function in *Drosophila*. Genetics *207,* 389-412.

Keller, R.B., Tran, T.T., Pyott, S.M., Pepin, M.G., Savarirayan, R., McGillivray, G., Nickerson, D.A., Bamshad, M.J., and Byers, P.H. (2017). Monoallelic and biallelic CREB3L1 variant causes mild and severe osteogenesis imperfecta, respectively. Genet. Med.

Maruyama, R., Grevengoed, E., Stempniewicz, P., and Andrew, D.J. (2011). Genome-Wide Analysis Reveals a Major Role in Cell Fate Maintenance and an Unexpected Role in Endoreduplication for the *Drosophila* FoxA Gene *Fork Head*. PLoS ONE *6,* e20901.

Myat, M.M., and Andrew, D.J. (2000b). Organ shape in the Drosophila salivary gland is controlled by regulated, sequential internalization of the primordia. Development *127,* 679-691.

Myat, M.M., and Andrew, D.J. (2000a). Fork head prevents apoptosis and promotes cell shape change during formation of the Drosophila salivary glands. Development *127,* 4217-4226.

Myat, M.M., and Andrew, D.J. (2002). Epithelial Tube Morphology Is Determined by the Polarized Growth and Delivery of Apical Membrane. Cell *111,* 879-891.

Nitta, K.R., Jolma, A., Yin, Y., Morgunova, E., Kivioja, T., Akhtar, J., Hens, K., Toivonen, J., Deplancke, B., Furlong, E.E.M., and Taipale, J. (2015). Conservation of transcription

factor binding specificities across 600 million years of bilateria evolution. ELife *4,* e04837.

Noyes, M.B., Meng, X., Wakabayashi, A., Sinha, S., Brodsky, M.H., and Wolfe, S.A. (2008). A systematic characterization of factors that regulate Drosophila segmentation via a bacterial one-hybrid system. Nucleic Acids Res. *36,* 2547-2560.

Roper, K. (2012). Anisotropy of Crumbs and aPKC drives myosin cable assembly during tube formation. Dev. Cell. *23,* 939-953.

Saito, A., Hino, S., Murakami, T., Kanemoto, S., Kondo, S., Saitoh, M., Nishimura, R., Yoneda, T., Furuichi, T., Ikegawa, S.*, et al.* (2009). Regulation of endoplasmic reticulum stress response by a BBF2H7-mediated Sec23a pathway is essential for chondrogenesis. Nat. Cell Biol. *11,* 1197-1204.

Sanchez-Corrales, Y., Blanchard, G.B., and Röper, K. (2018). Radially-patterned cell behaviours during tube budding from an epithelium. ELife *7,* e35717.

Shin, J.J.H., Gillingham, A.K., Begum, F., Chadwick, J., and Munro, S. (2017). TBC1D23 is a bridging factor for endosomal vesicle capture by golgins at the trans-Golgi. Nature Cell Biology

Weigel, D., Bellen, H.J., Jurgens, G., and Jackle, H. (1989). Primordium specific requirement of the homeotic gene fork head in the developing gut of the Drosophila embryo. Roux's Archives of Developmental Biology *198,* 201-210.

Yamamoto, S., Jaiswal, M., Charng, W., Gambin, T., Karaca, E., Mirzaa, G., Wiszniewski, W., Sandoval, H., Haelterman, N., Xiong, B.*, et al.* (2014). A Drosophila Genetic Resource of Mutants to Study Mechanisms Underlying Human Genetic Diseases. Cell *159,* 200-214.

PAGE INTENTIONALLY TO BE BLANK

CHAPTER 1

CrebA directly regulates both regulators and effectors of secretory capacity

Abstract

CrebA/Creb3L-like proteins coordinately regulate expression of the protein machinery of secretory organelles and boost expression of secretory cargo. To gain insight into the mechanism by which these bZip transcription factors regulate target gene expression in a tissue-specific manner, we developed an assay to identify CrebA DNA binding sites in the developing embryonic salivary gland (SG) of Drosophila. We demonstrate that CrebA binding is linked to open chromatin and the presence of consensus binding sites previously discovered through *in vitro* assays. We show that CrebA functions as a transcriptional activator and that it is necessary and/or sufficient for the expression of ~80% of the tested SG expressed genes to which CrebA binds *in vivo*. We demonstrate the importance of the consensus sites in CrebA transcriptional activation by mutating the consensus sites in the context of the endogenous enhancers of two of the newly discovered CrebA transcriptional targets, Xbp1 and Tudor-SN (TSN). Finally, we describe secretory roles for three newly discovered direct CrebA target genes, two regulators – Xbp1 and TSN – and one downstream secretory effector –Tbc1. The mammalian orthologues of both CrebA and Xbp1 have been implicated in the unfolded protein response. Our findings suggest that the ER stress induced by drugs and excess unfolded proteins activates endogenous pathways that normally function to accommodate the increased secretory load of professional secretory cells. Altogether, these findings emphasize the role of the CrebA/Creb3L-like proteins in the direct regulation of secretory capacity and suggest that many/most of the additional direct targets of CrebA identified in this study function in secretion.

Introduction

Organs are uniquely specialized for their functions. The lungs, pancreas, and liver have not only unique architectures but also unique gene expression profiles. For example, the lungs have elaborate tubular arborizations that provide large surface areas for gas exchange. The specialized type II cells of the lungs produce and secrete high levels of surfactant proteins, which serve to protect the lung tissue from alveolar collapse, reduce surface tension, and increase compliance (Mendelson, 2000). The pancreas comprises large secretory acinar cells linked by small ductal tubes to either the main or accessory pancreatic ducts. The pancreas is functionally divided into both an endocrine and exocrine pancreas, which produce and secrete high levels of either hormones (insulin and glucagon) or digestive enzyme precursors (amylase, lipase, and proteases such as trypsinogen), respectively (Jennings et al., 2015). The liver has an unusual architecture, which allows close contact between liver cells and the arteries and veins that carry the blood coming from the digestive tract to be appropriately altered before its passage to the rest of the body (Ober and Lemaigre, 2018). Genes encoding xenobiotic receptors and detoxifying enzymes are upregulated in the liver because of its role in removing toxins from the bloodstream (Larigot et al., 2018). The liver is important in the processing and storage of nutrients, and thus also encodes the enzymes involved in, for example, the conversion of glycogen to glucose for release into the blood as a quick energy source.

Organ specific gene expression is controlled by a unique combination of transcription factors (TFs). In the lung, surfactant expression is controlled by Thyroid transcription factor 1 (TTF-1, also known as Nkx2.1) and Hepatic nuclear factor 3 (HNF-3) (Mendelson, 2000; Yang et al., 2018). The pancreatic program of gene expression is activated by Pancreatic and duodenal homeobox 1 (PDX1) and Pancreas transcription factor 1A (PTF1A). Working with another transcription factor – the Gli Family Zinc Finger 3 (GLIS3), Pdx1 also specifies the endocrine progenitors and beta cell precursors in the pancreas (ZeRuth et al., 2013). The MyoD family of transcription factors controls expression of hundreds of muscle specific genes, including not only

the proteins of the contractile unit of muscle, known as the sarcomere, but also regulators of mitochondria biogenesis such as PCG-1β, since muscle cells have very high energy requirements (Shintaku et al., 2016).

Many specialized organs have exaggerated needs for what are considered housekeeping functions, such as the increases in mitochondrial biogenesis observed in muscles. Professional secretory cells, like those found in the pancreas, salivary and mammary glands, must increase secretory capacity to produce the high levels of digestive enzymes, hormones, saliva and milk required for organismal homeostasis and/or offspring survival. The CrebA/Creb3-like family of TFs has been shown to expand secretory capacity (Abrams and Andrew, 2005; Barbosa et al., 2013; Fox et al., 2010). Within Drosophila, where the general role of CrebA/Creb3-like family members was first revealed, loss of the single family member CrebA results in significant decreases in the expression of secretory pathway component genes (SPCGs) in tissues where these components must be upregulated for the organ to function properly, such as the salivary gland (SG) and epidermis. CrebA larvae are unable to hatch from the eggshell and have little or no cuticle compared to wild type (Abrams and Andrew, 2005; Fox et al., 2010). This is due to the lack of secretion of cuticle components by the epidermis. The upregulation of the SPCGs, which is highest in the SG where CrebA expression is highest and where its function was first characterized, has been shown to be direct for a small subset of SPCGs (Fox et al., 2010). Mice null for two CrebA orthologues, Creb3L1 [aka OASIS] and Creb3L2 [aka BBF2H7], have defects in osteogenesis and chondrogenesis, respectively (Saito, 2014). Both phenotypes are consistent with the Creb3L1 and Creb3L2 proteins functioning to increase secretion in the respective cell types (osteoblasts/osteocytes and chondroblasts/chondrocytes). Additionally, members of a family carrying a point mutation in the Creb3L1 gene develop osteogenesis imperfecta, also known as brittle bone disease (Keller et al., 2017). Ectopic expression of CrebA, as well as ectopic expression of the active form of any of the five human Creb3-like transcription factors, in the Drosophila embryo leads to upregulation of all SPCGs that have been tested to date (Barbosa

et al., 2013; Fox et al., 2010). Altogether, the phenotypes associated with loss of *CrebA* and its closest human orthologues, as well as the activity of the protein in Drosophila embryos, support a universal role for CrebA/Creb3-like proteins in the upregulation of secretory machinery.

Interestingly, whereas CrebA regulates expression of the protein components of the secretory organelles, including the endoplasmic reticulum (ER), Golgi, transport and secretory vesicles, it does not upregulate lipid biogenesis (Fox et al., 2010). As a consequence, when *CrebA* function is lost, there is an increase in plasma membrane accumulation at the septate junctions (the invertebrate equivalent to tight junctions), which presumably represents the excess lipid that would normally populate the ER, Golgi and secretory vesicle membranes (Fox and Andrew, 2015).

Microarray studies in developing embryos comparing *CrebA* null mRNA levels to WT levels revealed not only decreases in SPCG transcripts, but also decreases in transcript levels for genes encoding secretory cargo (Fox et al., 2010). It is unclear whether regulation of cargo genes by CrebA is direct or indirect since CrebA is required to achieve full wild-type levels of expression of Sage, an SG-specific bHLH transcription factor that works with the Drosophila FoxA orthologue, known as Fkh, to active SG expression of secreted cargo.

To find additional CrebA targets and to distinguish between direct and indirect transcriptional regulation, chromatin precipitation with a CrebA antibody followed by sequencing of the DNA that precipitates with the CrebA protein (ChIP-Seq) would be a good approach. Since CrebA is expressed in multiple embryonic tissues and since the embryonic SG is too small to surgically isolate, it is difficult to use this approach to identify SG-specific binding of CrebA in the embryo. To circumnavigate these challenges, we use a variation on ChIP-Seq that allows us to identify SG-specific binding events. Specifically, we perform ChIP-seq using GFP antibodies on chromatin isolated from embryos expressing a GFP-tagged and functional form of CrebA in the SG. For quality control, we compare the binding events identified using this strategy to those observed with chromatin isolated from whole embryos and immuno-precipitated with CrebA

antiserum. Our experiments reveal that CrebA binds a large portion of the genome (~1%) and that CrebA is binding to open chromatin. Our analysis of the CrebA-bound DNA suggests that binding is driven by the occurrence of consensus CrebA binding motifs within this open chromatin. We demonstrate that CrebA is necessary and/or sufficient for wild-type levels of SG expression for a large subset of SG genes bound by CrebA but not previously known to be CrebA targets. Three of the genes for which CrebA function is both necessary and sufficient for their expression are *Xbp1*, *TSN*, and *tbc1*, which we implicate in secretory function. Using CRISPR/Cas9-directed mutagenesis, we show that two key consensus CrebA binding sites in the *Xbp1* gene and one in the *TSN* gene are necessary for their CrebA-dependent expression and their secretory function in two different tissues.

Results

Approach to identifying tissue-specific CrebA binding sites

CrebA is expressed in multiple embryonic tissues, including the salivary gland (SG), proventriculus, trachea, and epidermis (Fig 2A). Thus, to identify *in vivo* CrebA binding sites specifically in the SG, we used the Gal4/UAS system to express a *CrebA* transgene tagged at its C-terminus with GFP (CrebA-GFP) for ChIP-Seq (chromatin immunoprecipitation followed by DNA sequencing) using a well-characterized GFP antibody that has been used extensively for ChIP-Seq studies of Drosophila transcription factors (Kittler et al., 2013; Kudron et al., 2018; Loganathan et al., 2016; Negre et al., 2011). Two Gal4 drivers were used, *fkh-Gal4* and *sage-Gal4*; both drive UAS-transgene expression in multiple tissues, but have only SG expression in common (Fig 2B-C).

To test the functionality of the *UAS-CrebA-GFP* transgene, we expressed both untagged *UAS-CrebA* and *UAS-CrebA-GFP* in embryonic stripes using the *en-Gal4* driver and examined expression of *Spase12,* a known downstream target of CrebA (Fig 3A,B; (Fox et al., 2010). Expression of both transgenes resulted in ectopic expression of *Spase12* mRNA in stripes (Fig 3B), with levels roughly corresponding to the levels of detectable CrebA protein (Fig 3A). We also expressed both *UAS-CrebA* and *UAS-CrebA-GFP* in alternating segments using *prd-Gal4* in the background of both wild-type and *CrebA* null animals, and examined the resulting first instar larval cuticles. Consistent with the expression domain of the *prd*-Gal4 line (Fig 3C,D), expression of both *CrebA* constructs resulted in changes in dorsal hair morphology in approximately every other segment in otherwise WT larvae (Fig 3E-G"). Loss of *CrebA* lead to a loss of cuticular structures on all larval surfaces, with barely visible denticles, as previously reported (Fig 3H; Abrams and Andrew, 2005; Fox et al., 2010). Expression of either untagged or GFP-tagged CrebA in a *CrebA* mutant rescued the cuticle defects in the alternating segments where *prd-Gal4* is expressed (Fig 3I-J). Finally, we immunostained SG polytene chromosomes expressing the UAS-CrebA-GFP transgene under the control of *fkh*-Gal4 with CrebA antiserum, which

Figure 2: Experimental approach to identify CrebA binding sites in whole embryos and salivary glands (SG). (A) CrebA expression was detected with the rabbit polyclonal CrebA antiserum used for whole embryo ChIP-Seq, shown at embryonic stages 10, 14, and 16 (ventral views). Black arrowhead: SG; white arrowhead: proventriculus; black arrows: epidermis; white arrows: trachea. (B-C) GFP expression was observed in the SG with *UAS-CrebA-GFP* expressed under the control of either *fkh-Gal4* (B) or *sage-Gal4* (C) at embryonic stages 10, 14, and 16 (ventral views st10 and st16, lateral views st14). Black arrowheads: SG. (D) The cartoon outlines the scheme to identify CrebA binding sites in whole embryos and in SGs using ChIP-Seq. (E) Salivary gland polytene chromosomes from WT larvae (top) and larvae expressing CrebA-GFP under the control of fkh-Gal4 (bottom). Note that the bands on wild-type chromosomes stained with CrebA antiserum made in either rat or rabbit completely overlap. Similarly, the bands that stain with CrebA (rat) antiserum completely overlap the bands that stain with GFP antiserum (rabbit).

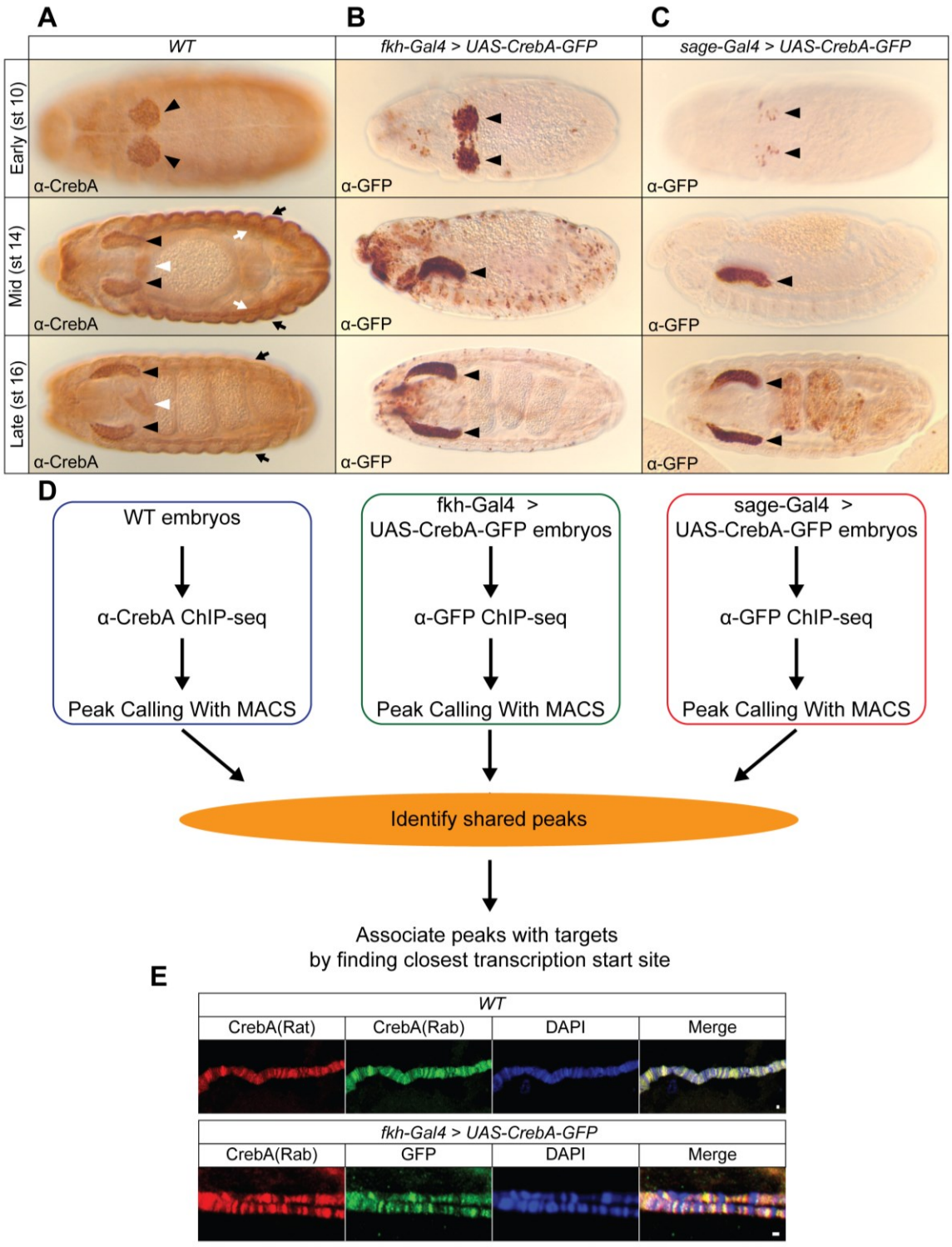
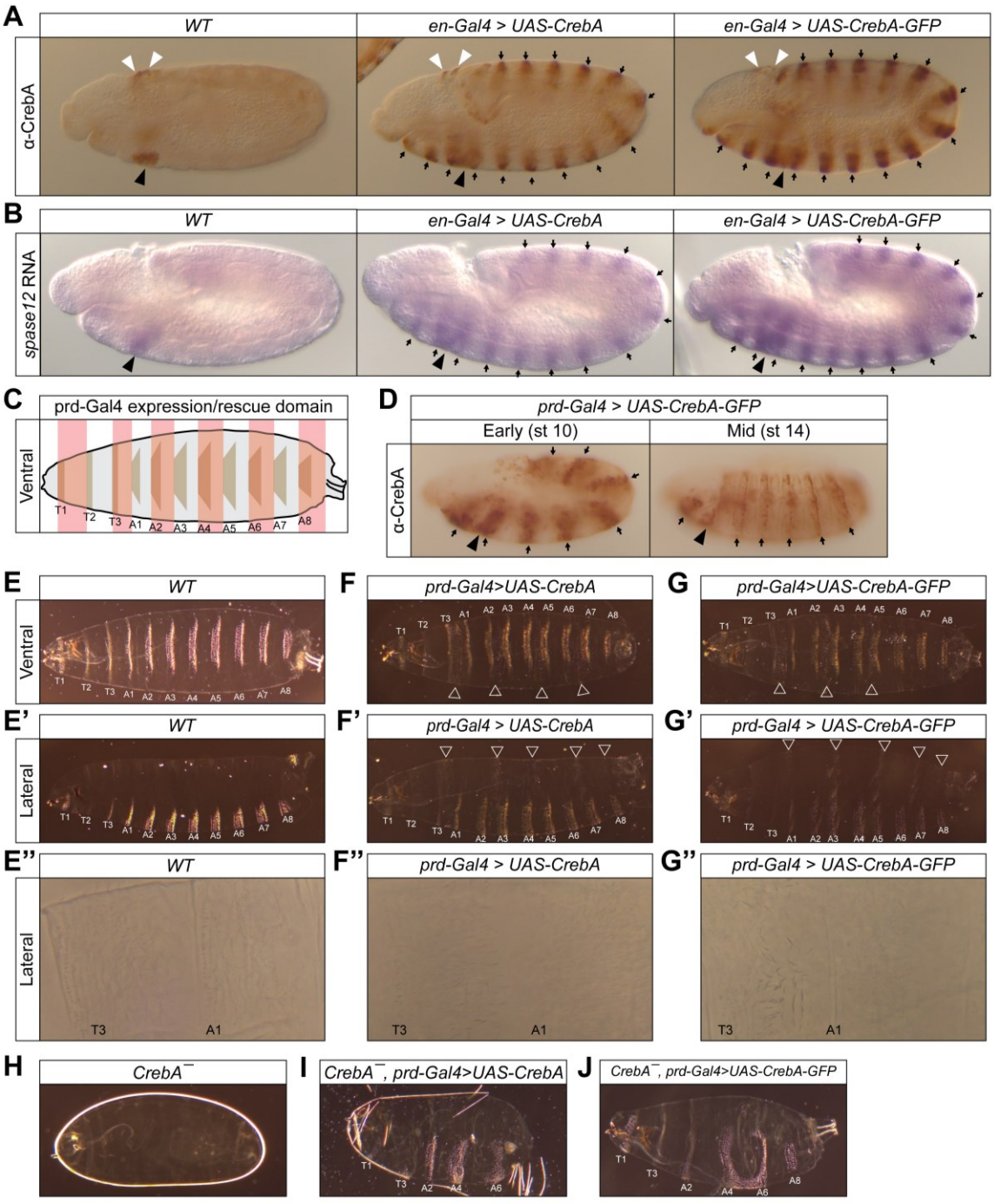

A
B
C
WT
fkh-Gal4 > UAS-CrebA-GFP
sage-Gal4 > UAS-CrebA-GFP
Early (st 10)
Mid (st 14)
Late (st 16)
α-CrebA
α-GFP
D
WT embryos
α-CrebA ChIP-seq
Peak Calling With MACS
fkh-Gal4 > UAS-CrebA-GFP embryos
α-GFP ChIP-seq
Peak Calling With MACS
sage-Gal4 > UAS-CrebA-GFP embryos
α-GFP ChIP-seq
Peak Calling With MACS
Identify shared peaks
Associate peaks with targets by finding closest transcription start site
E
WT
CrebA(Rat)
CrebA(Rab)
DAPI
Merge
fkh-Gal4 > UAS-CrebA-GFP
CrebA(Rab)
GFP
DAPI
Merge

Figure 3: GFP-tagged CrebA tagged is fully functional. (A) WT embryo (left), and embryos expressing either untagged CrebA (middle) or GFP-tagged CrebA (right) under the control of *en-Gal4* and stained with CrebA antiserum. Black arrows: ectopic CrebA expression in segmental stripes; black arrowheads: endogenous CrebA expression in the SG; white arrowheads: endogenous CrebA expression in amnioserosal cells. (B) *in situ* hybridizations with a probe to a known target of CrebA, *spase12*. Shown are a WT embryo (left), and embryos expressing either untagged CrebA (middle) or GFP-tagged CrebA (right) under the control of *en-Gal4*. Black arrows: ectopic *spase12* expression in segmental stripes; black arrowheads: endogenous *spase12* expression in the SG. (C) Cartoon of *prd-Gal4* expression pattern in larvae. (D) CrebA protein in embryos expressing *UAS-CrebA-GFP* under control of *prd-Gal4*. (E-J) Cuticle preps of 1st instar larvae. (E) Ventral view of a WT larva (E) and larvae expressing CrebA transgenes under the control of *prd-Gal4* (F,G). (E'–G') Lateral view of a WT larva (E') and larvae expressing CrebA transgenes under the control of *prd-Gal4* (F',G'). (E" – G") Higher magnification phase images of lateral views of a WT larva (E") and larvae expressing CrebA transgenes under the control of *prd-Gal4* (F",G"). Arrowheads indicate regions with thicker hair structures on the dorsal surface in cells expressing either untagged or tagged CrebA. (H) *CrebA* mutant larval cuticle. (I) Region-specific rescue of *CrebA* mutant by untagged CrebA driven by *prd-Gal4* in alternating segments. (J) Region-specific rescue of *CrebA* mutant by GFP-tagged CrebA driven by *prd-Gal4* in alternating segments.


A
WT
en-Gal4 > UAS-CrebA
en-Gal4 > UAS-CrebA-GFP
α-CrebA
B
WT
en-Gal4 > UAS-CrebA
en-Gal4 > UAS-CrebA-GFP
spase12 RNA
C
prd-Gal4 expression/rescue domain
Ventral
T1 T2 T3 A1 A2 A3 A4 A5 A6 A7 A8
D
prd-Gal4 > UAS-CrebA-GFP
Early (st 10)
Mid (st 14)
α-CrebA
E
WT
Ventral
F
prd-Gal4>UAS-CrebA
G
prd-Gal4>UAS-CrebA-GFP
E'
WT
Lateral
F'
prd-Gal4 > UAS-CrebA
G'
prd-Gal4 > UAS-CrebA-GFP
E''
WT
Lateral
T3 A1
F''
prd-Gal4 > UAS-CrebA
T3 A1
G''
prd-Gal4 > UAS-CrebA-GFP
T3 A1
H
CrebA⁻
I
CrebA⁻, prd-Gal4>UAS-CrebA
J
CrebA⁻, prd-Gal4>UAS-CrebA-GFP

should detect both endogenous and expressed GFP-tagged CrebA, and anti-GFP, which should detect only the expressed GFP-tagged CrebA. Antisera recognizing both proteins completely colocalized (Fig. 2E). Altogether, these findings indicate that the GFP-tagged CrebA protein is functionally equivalent to the untagged protein both in its ability to activate target gene expression in new domains, in its ability to rescue loss of *CrebA* phenotypes, and in its ability to bind SG chromatin.

Identification of SG-specific CrebA binding sites

In vivo CrebA binding sites were identified by comparing independent biological replicates of all three experiments to input DNA controls (from non-immunoprecipitated chromatin) (Fig. 2D). A large number of binding peaks, called by the MACS program (Zhang et al., 2008), were observed in all three experiments (Fig. 5A): 14,012 peaks were observed with WT chromatin pulled down with αCrebA (referred to as the WT-CrebA set); 6,457 peaks were observed with *fkh-Gal4::UAS-CrebA-GFP* chromatin precipitated with αGFP (referred to as the fkh-GFP set); and 7,903 peaks were observed with *sage-Gal4::UAS-CrebA-GFP* chromatin immunoprecipitated with αGFP (referred to as the sage-GFP set). We used the Intersect function through the Galaxy platform (usegalaxy.org) to identify approximately 4000 sites common to all three experiments, with the exact number depending on the order of the comparisons (Fig 4); 3884 binding sites were identified when the comparisons began with the WT-CrebA set and either of the GFP sets and 4303 binding sites were identified when the comparisons were done with the GFP sets first (Fig 5A, Fig 6A,A').

Between 80 and 88% of the binding intervals common to the sage-GFP and fkh-GFP sets were also observed in the WT-CrebA set, indicating that the approach to find SG-specific CrebA binding sites worked well. Between 28% and 31% of sites in the WT-CrebA set were also in the SG-specific set, indicating that CrebA does not bind the same sequences in all tissues and that sites bound in the SG represent a substantial part of the total. The intervals shared between

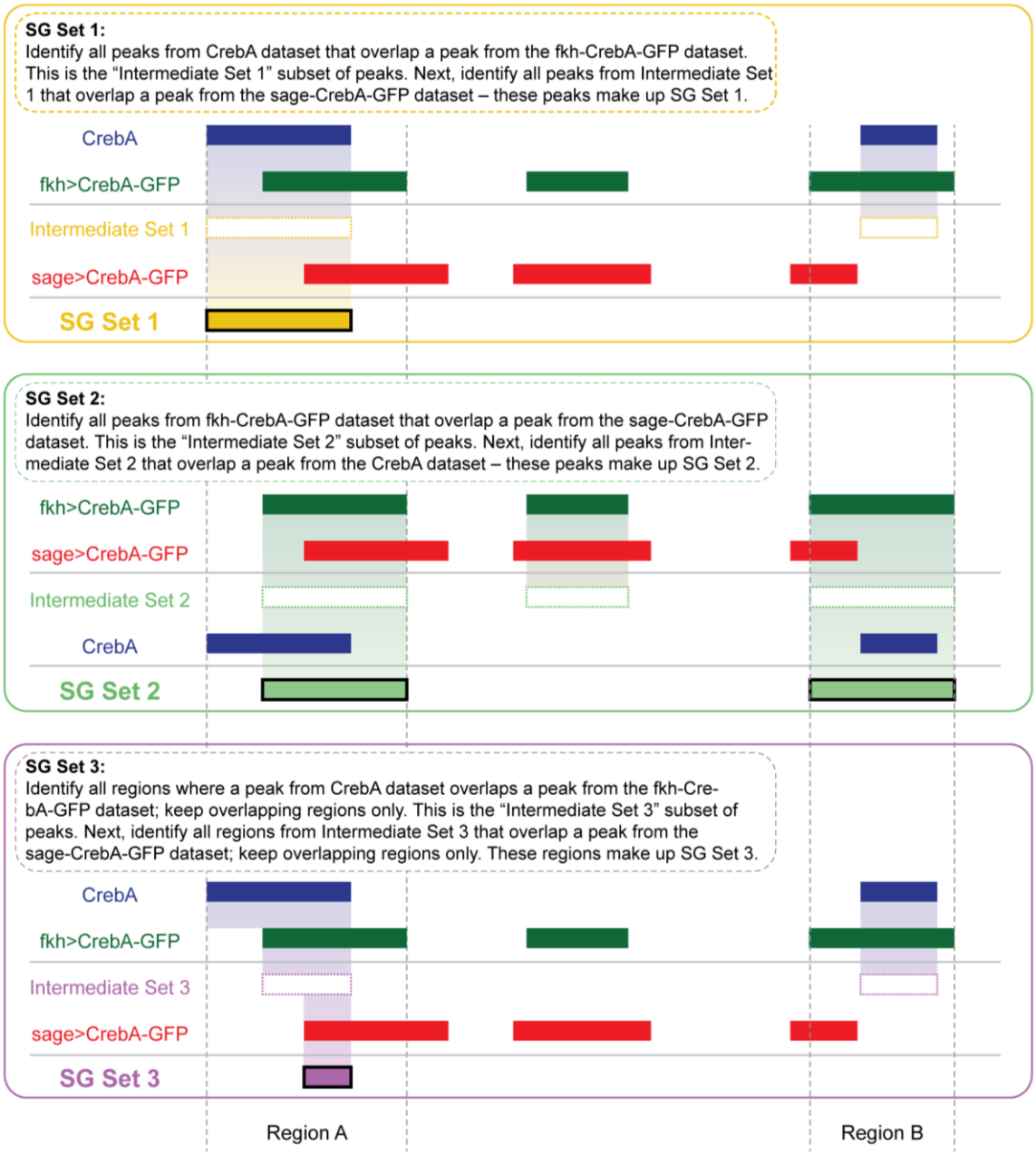


Figure 4: Diagram of how each dataset was generated by using Galaxy intersection function.

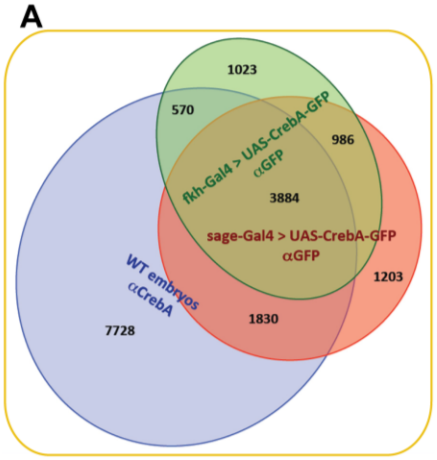
the fkh-GFP and sage-GFP sets but not found in the WT-CrebA set suggest that either the increased level of CrebA expression is driving increased occupancy or that the GFP pull-downs are more sensitive. For simplicity, we will refer to the sets of genes pulled down by all three experiments as the SG-specific sets. SG-specific set 1 comprises the 3884 sites obtained in the intersection starting with the WT-CrebA data (Fig. 5A) and SG-specific set 2 comprises the 4303 sites obtained in the intersection starting with the fkh-GFP and sage-GFP data (Fig 6A).

We also compared the three ChIP-Seq datasets using a variation of the Galaxy program that identifies the exact interval of base pair overlap in the data from all three experiments. This analysis identified the same 5115 sequences regardless of the order of comparisons (Fig 6A'). Notably, 90.8% of the sequences that overlap between the fkh-GFP and sage-GFP sets were also present in the WT-CrebA set, again indicating that the experimental approach worked. We refer to this set of sequences as SG-specific set 3.

To gain insight into what drives CrebA SG chromatin binding, we ran each of the SG-specific sets through the SepPos motif analysis program in the Cistrome/Galaxy Platform that searches for sequences matching the motifs of known transcription factors (Liu et al., 2011). The number one sequence cluster identified from the analysis of all three SG-specific sets corresponded to the consensus binding sites for Creb3L1 and Creb3, two known mammalian orthologues to Drosophila CrebA (Fig. 5B, 6B,B'). This cluster also includes Xbp1, a transcription factor that plays a major role in the unfolded protein response (UPR) and that binds to approximately the same consensus sequence. The second sequence cluster discovered for SG-specific set 1 was Creb3L2 (Fig. 5B), another CrebA mammalian orthologue, which is also found as cluster 4 in SG-specific sets 2 and 3 (Fig 6B,B'). Otherwise, the remaining top clusters for SG-specific set 1 include a motif for FoxN1 (Fig. 5B) and a variety of different transcription factors (not shown), whereas motifs for Helix-Loop-Helix (HLH) family members were prominent in the analysis of SG-specific sets 2 and 3 (Fig. 6B.B'). The differences in motifs identified among the three sets likely reflect the biases in the overlap analysis that produced them – SG-specific set 1 is

Figure 5: CrebA ChIP-seq binding and motif enrichment for Set 1. (A) Venn diagram of ChIP-seq peaks and overlaps. (B) SepPos analysis of motifs enriched in the intersection of all three datasets identified consensus binding sites for several known mammalian transcription factors. The motif logo for the transcription factors in the first cluster is shown below the cluster table. (C) Percentage of binding peaks with at least one identifiable CrebA consensus binding site, based on FIMO analysis of all CrebA motifs either individually or combined. The CrebA consensus binding sites, obtained from previously published studies are shown. (D) Centrimo analysis of 250bp on either side of peak summits from the intersection of all three datasets; summits were determined using the fkh-Gal4 driven dataset. All three of the consensus motifs analyzed show enrichment around the summit. (E) The top 10 categories of GO terms based on DAVID analysis of called targets that have peaks in all three datasets.

A



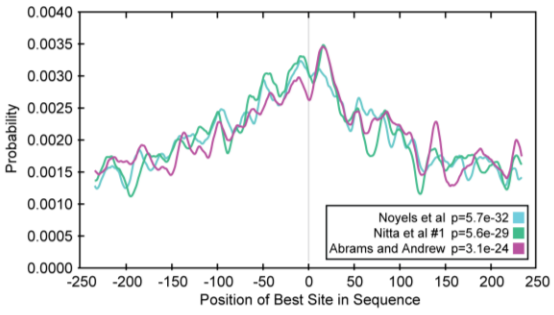
B

Clusters	Motif ID	TF	DNA Binding Domain	Hits	-10*log(pval)
1	MS00304	CREB3L1	Leucine Zipper Family	3883	213.543
	MS00308	CREB3	Leucine Zipper Family	3630	98.33
	MS00342	XBP1	Leucine Zipper Family	752	85.453
2	MS00311	Creb3l2	Leucine Zipper Family	3327	112.685
3	M00332	FOXN1	Forkhead Domain Family	2687	95.642

C

	SG-specific Set 1	SG-specific Set 2	SG-specific Set 3
Abrams and Andrew	77.7%	62.7%	50.8%
Nitta et al #1	82.7%	67.7%	56.4%
Nitta et al #2	60.7%	44.4%	34.4%
Nitta et al #3	67.9%	52.5%	41.5%
Noyes et al	78.9%	64.0%	52.5%
Any of the five motifs	94.5%	85.8%	75.4%

D

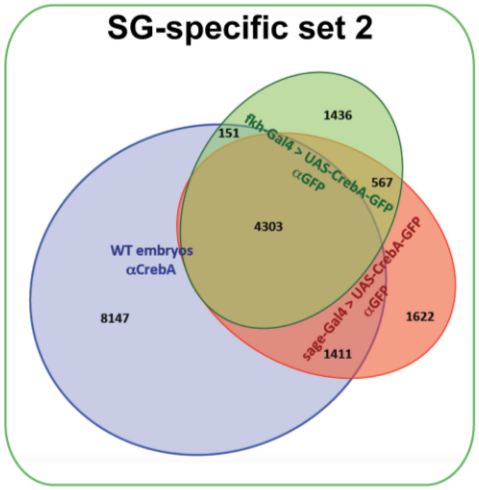


E

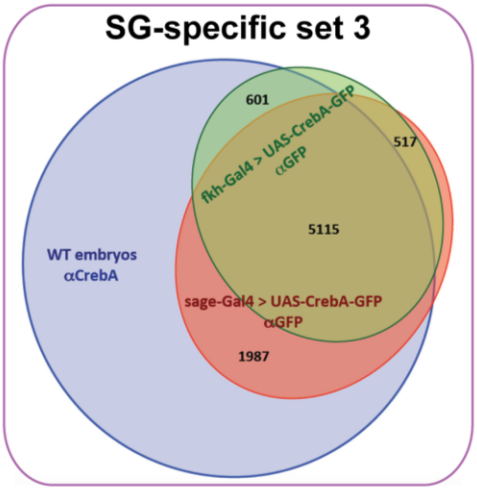
DAVID Category	Enrichment	p-Value	Benjamini
Transcription	5.94	1.9E-09	1.0E-07
Nucleotide Binding	4.23	3.0E-06	1.1E-04
Zinc Finger	3.35	8.4E-07	3.5E-05
Transcription Factor	3.06	5.5E-06	3.0E-03
PH Domain	2.36	1.6E-05	4.0E-02
GTP Binding	2.29	4.8E-05	1.8E-02
BTB Domain	2.11	5.7E-03	7.4E-01
Protein Deubiquitination	1.93	3.7E-03	9.1E-01
Septate Junction	1.87	8.2E-06	8.5E-04
TCA Cycle	1.85	1.9E-03	7.7E-02

Figure 6. Parallel analyses of ChIP-seq binding site data for SG sets 2 and 3. (A,A') Proportional Venn Diagrams showing the overlap in binding sites from all three experiments. (B, B') Identification of top motifs for known mammalian transcription factors enriched in the intersection of all three datasets. (C,C') Centrimo analysis of 250bp on either side of peak summits from the intersection of all three datasets; summits were determined using the fkh-Gal4 driven dataset. All three of the consensus motifs analyzed show enrichment around the summit. (D,D') The top 10 categories of GO terms based on DAVID analysis of called targets that have peaks in all three datasets.

A



A'



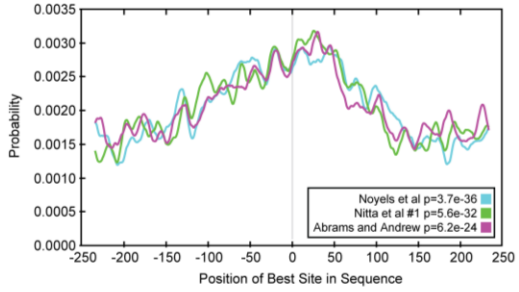
B

Clusters	Motif ID	TF	DNA Binding Domain	Hits	-10*log(pval)
1	MS00304	CREB3L1	Leucine Zipper Family	3926	265.456
	MS00342	XBP1	Leucine Zipper Family	4187	180.131
	MS00308	CREB3	Leucine Zipper Family	2360	159.734
2	MC00254	E2F6	Transcription Factor Family	3375	253.907
3	MC00124	MAX	Helix-Loop-Helix Family	3332	248.4
	MC00461	Max	Helix-Loop-Helix Family	2667	233.934
	MS00260	HEY2	Helix-Loop-Helix Family	3110	231.975
	MC00394	MYC	Helix-Loop-Helix Family	4295	189.024
	MC00532	Mycn	Helix-Loop-Helix Family	4144	164.311
	MS00258	HEY1	Helix-Loop-Helix Family	3122	161.064
	MC00373	Mxi1	Helix-Loop-Helix Family	3882	154.086
	MS00267	MNT	Helix-Loop-Helix Family	2933	111.949
	MC00407	Myc	Helix-Loop-Helix Family	2948	109.562
	MS00253	CLOCK	Helix-Loop-Helix Family	2940	89.219

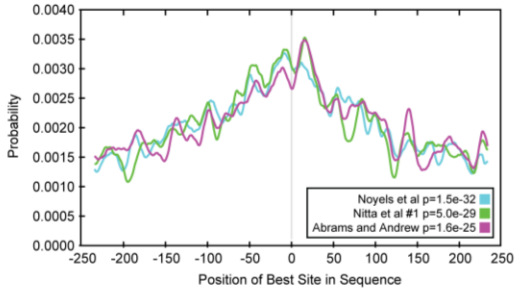
B'

Clusters	Motif ID	TF	DNA Binding Domain	Hits	-10*log(pval)
1	MS00304	CREB3L1	Leucine Zipper Family	3260	654.878
	MS00342	XBP1	Leucine Zipper Family	3448	491.293
	MS00308	CREB3	Leucine Zipper Family	3482	490.331
2	MS00258	HEY1	Helix-Loop-Helix Family	4869	638.129
	MC00461	Max	Helix-Loop-Helix Family	4921	575.43
	MC00394	MYC	Helix-Loop-Helix Family	4981	507.823
	MC00124	MAX	Helix-Loop-Helix Family	3368	501.001
	MC00373	Mxi1	Helix-Loop-Helix Family	4968	495.199
	MC00407	Myc	Helix-Loop-Helix Family	3470	429.176
	MC00532	Mycn	Helix-Loop-Helix Family	4920	423.782
	MS00260	HEY2	Helix-Loop-Helix Family	4636	423.033
	MS00267	MNT	Helix-Loop-Helix Family	4450	363.494
	MC00362	Arntl	Helix-Loop-Helix Family	4363	310.713
	MS00253	CLOCK	Helix-Loop-Helix Family	3605	102.162
3	MC00224	EPAS1	Helix-Loop-Helix Family	4071	589.169
	MC00316	HIF1A	Helix-Loop-Helix Family	4985	504.911
	MC00262	ARNT	Helix-Loop-Helix Family	2856	414.93

C



C'



D

DAVID Category	Enrichment	p-Value	Benjamini
Transcription	4.7	2.7E-07	1.5E-05
Nucleotide Binding	3.48	4.9E-05	1.4E-03
Transcription Factor	3.19	6.2E-06	3.4E-03
Zinc Finger	2.6	4.3E-06	1.8E-04
PH Domain	2.3	3.1E-05	7.7E-02
Ribosomal Protein	2.12	1.4E-05	5.1E-04
MADF Domain	2.07	1.7E-03	7.6E-01
Septate Junction	2.06	7.9E-06	7.0E-04
Protein Deubiquitination	1.96	3.2E-03	1.1E-01
Protein Export (SPCG)	1.82	1.3E-04	7.7E-03

D'

DAVID Category	Enrichment	p-Value	Benjamini
Transcription	4.62	3.6E-07	2.0E-05
Nucleotide Binding	3.63	1.9E-05	6.2E-04
Zinc Finger	3.26	1.4E-06	5.6E-05
Transcription Factor	2.92	1.1E-05	6.2E-03
PH Domain	2.45	1.5E-05	3.8E-02
Chaperone	2.05	2.6E-03	3.8E-01
Protein Deubiquitination	1.99	2.9E-03	9.1E-01
BTB Domain	1.96	8.9E-03	6.5E-01
Septate Junction	1.94	5.5E-06	5.7E-04
GTP Binding	1.92	1.7E-04	5.9E-02

biased toward CrebA binding in whole embryos and could reflect binding sites for the range of tissue-specific factors that may cooperate with CrebA for gene regulation and/or differences in chromatin accessibility in the different tissues. SG-specific sets 2 and 3 are biased toward CrebA binding in the SG and could reflect SG-specific transcription factors and/or chromatin accessibility. Notably, an HLH transcription factor known as Sage, which is expressed to very high levels in embryonic SGs, has previously been implicated in controlling SG-specific gene expression (Fox et al., 2013).

We used FIMO (Bailey et al., 2009) to explore the extent to which *in vitro* CrebA consensus motifs (Abrams and Andrew, 2005; Fox et al., 2010; Nitta et al., 2015; Noyes et al., 2008) can account for *in vivo* CrebA SG binding. Notably, 94.5% of the bound sequences in SG-specific set 1 contain at least one CrebA consensus-binding site (p-value = 0.001) (Fig. 5C). Bound sequences in SG-specific sets 2 and 3 contained at least one identifiable CrebA consensus sequence 85.8% and 74.5% of the time, respectively. The somewhat smaller number in SG-specific set 3 may reflect how this set was generated, where only the exact (partial) sequences found in all three experiments were included.

Centrimo analysis was done to determine where in the binding peaks the most represented CrebA consensus sites are located (Bailey et al., 2009). This analysis of 500 nt regions revealed that the consensus sites tend to be close to the summit of each peak (Fig. 5D; Fig. 6C,C'). Thus, CrebA consensus sites, determined through *in vitro* DNA binding studies, are driving the *in vivo* binding of CrebA as revealed by ChIP seq.

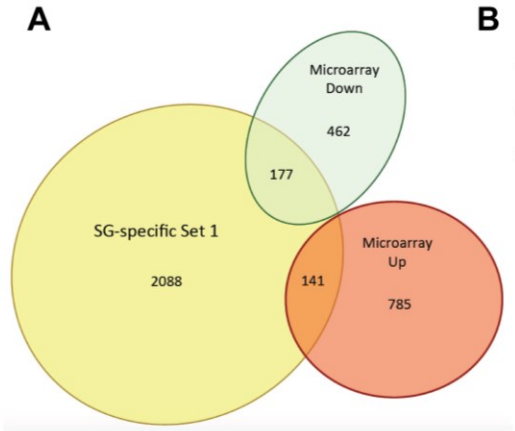
To identify the genes associated with the SG-specific sets of binding sites, we assigned each binding site to a target gene based on the nearest protein coding gene transcription start site (TSS) (Slattery et al., 2011). There are 2406 associated protein coding genes that have peaks in SG-specific set 1, 2397 in SG-specific set 2, and 2339 in set 3; 2236 protein coding genes (87.1%) are shared among all three sets. DAVID gene ontology analysis was done to learn what classes of genes CrebA binds in the SG (Huang da et al., 2009a; Huang da et al., 2009b).

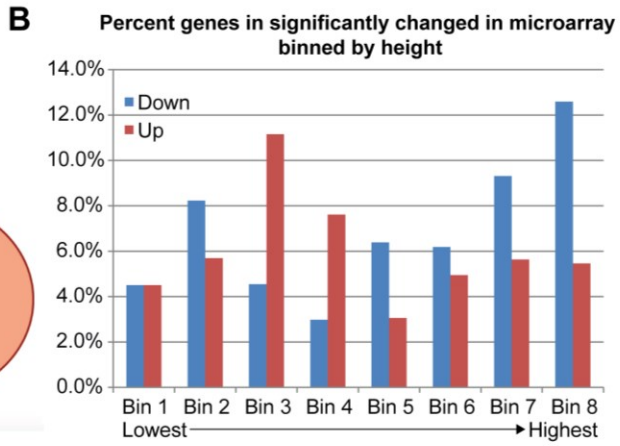
Regardless of which SG-specific set was used for this analysis, similar types of functional clusters were revealed (Fig. 5E; Suppl. Fig. 6D,D'). Most notably, transcription factors and other nucleic acid binding proteins emerge in the top categories with all searches. Other prominent clusters were PH-domain containing proteins, septate junctions (SJs) and GTPases. These findings were surprising, since previous expression studies indicate that CrebA primarily regulates secretory pathway component genes and secretory cargo (Fox et al., 2010). This outcome suggests that CrebA binding per se does not indicate gene regulation and may instead be simply a consequence of chromatin accessibility combined with the presence of one or more CrebA consensus-binding site.

SG binding sites and links to gene regulation

To begin to explore how CrebA binding is linked to CrebA-dependent gene expression, we compared genes identified by the CrebA ChIP-Seq experiments to those whose expression is known to change in *CrebA* mutants based on previous microarray analyses (Fox et al., 2010). This comparison was done for each SG-specific class; in the best scenario, represented by the SG-specific set 1 data, 15.2% of genes that go up and 27.7% of genes that go down in *CrebA* mutants relative to WT are bound by CrebA in the SG (Fig 7A). To ask if increased site occupancy correlates with gene regulation, we binned binding sites according to peak height and then plotted peak height against the microarray genes known to go up or down in *CrebA* mutants (Fig 7B). Consistent with CrebA binding to a higher percentage of genes that go down in *CrebA* mutants, we see a trend of increased binding and increased percentage of genes that require CrebA for their expression. No such trend was observed with genes whose expression went up in the absence of *CrebA*, suggesting that CrebA primarily functions as a transcriptional activator. DAVID analysis of the categories of genes that are both bound by CrebA and whose expression goes down in *CrebA* mutants relative to WT based on the microarray studies revealed Secretory Pathway Component Genes (SPCGs) in the top four categories (Fig. 7C).

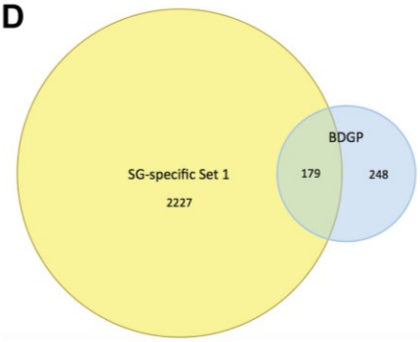
Figure 7. CrebA ChIP-Seq binding and CrebA regulation. (A) Proportional Venn diagram of overlap in target genes bound by CrebA in all three experiments and genes that go significantly up or down in microarray experiments comparing RNA from WT embryos and embryos null for *CrebA*. (B) Percentage of *CrebA*-dependent genes identified in microarray studies plotted relative to the strength of binding based ChIP-Seq analysis. (C) The top five categories of GO terms based on DAVID analysis of the overlap of called targets from CrebA binding from ChIP-Seq and genes whose expression goes down in *CrebA* mutants based on microarray studies. (D) Overlap in called targets from CrebA binding from ChIP-Seq and genes identified by BDGP as being expressed in the embryonic SG. (E, E') Super-resolution images of SG polytene chromosomes stained with CrebA antiserum (red) and signal quantification reveal that CrebA binds uncondensed chromatin, which does not stain intensely with DAPI (blue).

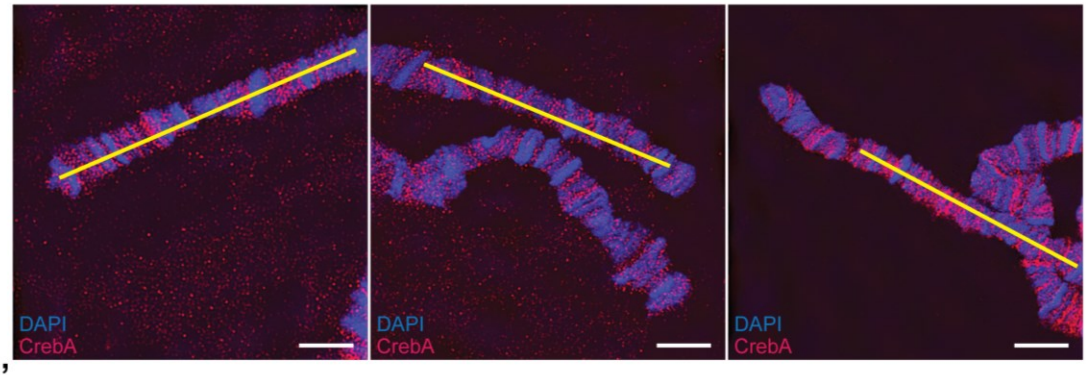


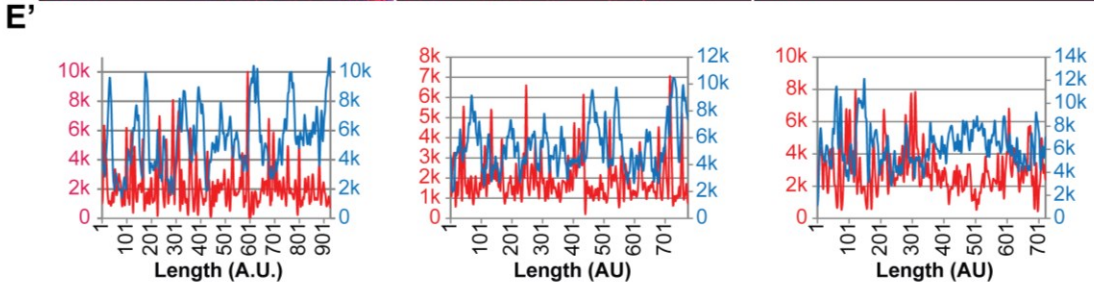


C

DAVID Category	Enrichment	p-Value	Benjamini
Transmembrane (SPCG)	6.13	5.7E-08	3.4E-06
SRP (SPCG)	5.55	8.1E-08	3.7E-06
Golgi Vesicle Transport (SPCG)	3.33	1.6E-04	1.0E-02
Signal Peptide Processing (SPCG)	2.05	1.2E-04	8.4E-03
DnaJ Domain	1.64	1.2E-04	3.3E-01







To test the idea that CrebA binding is linked to permissive chromatin, we first asked what percentage of genes expressed in the embryonic SG are bound by CrebA, since the chromatin surrounding these genes should be open in the SG during at least some embryonic stages. Of the 427 genes identified as SG-expressed by the Berkeley Drosophila Genomic Project (BDGP), CrebA bound 179 (42%) (Fig. 7D); this finding suggests that CrebA binding is linked to active chromatin in the SGs. Open chromatin in any cell type is estimated to be around 20% of the Drosophila genome (Milon et al., 2014).

To ask more directly if CrebA binds open versus condensed chromatin in SG cells, we stained third instar larval SG polytene chromosomes with CrebA and DAPI and imaged using super-resolution microscopy. We found that CrebA binding in the SG is to regions of open chromatin, i.e. regions of low intensity DAPI staining (Fig. 7E,E'). Since between 74.5 and 94.5% percent of sites bound in the embryonic SGs contain at least one CrebA consensus site, these findings indicate that CrebA binding occurs in open chromatin that contains at least one of the consensus sites previously defined in published *in vitro* binding studies.

CrebA activates expression of its direct SG targets *in situ*

To explore the relationship between CrebA SG binding and CrebA gene regulation, we did whole mount *in situ* hybridizations with probes specific to 33 SG-expressed genes (BDGP) in WT embryos, *CrebA* null embryos, and embryos expressing either GFP-tagged or untagged CrebA in en-Gal4 driven stripes (Table 1). The 33 SG genes included eight that did not show CrebA binding, three that were bound in the CrebA-GFP (SG) experiments only, and 22 that were bound by CrebA in all three sets of ChIP-Seq experiments. The 33 genes included nine that encode transcription factors (Table 1, green shading), between 11-13 that encode predicted secretory pathway component genes (SPCGs; Table 1, blue shading), and between 11-13 that encode a variety of other types of proteins.

Gene name	Fbgn	cDNA	CrebA Bound	Microarray Fold Change	CrebA Necessary	CrebA Sufficient	Known or predicted protein function/localization	CrebA binding sites
Myc	FBgn0262656	LD32539	No	n.s.	No	No	Transcription Factor	
Aatf	FBgn0031851	RE30678	No	-1.30	Yes	No	Transcription Factor	
CG5728	FBgn0039182	LD41803	No	n.s.	No	No	Alternative Splicing Regulator	
qsm	FBgn0028622	GH08941	No	n.s.	No	No	Circadian Rythm, GPI Anchored	
Tsp96F	FBgn0027865	LD19727	No	n.s.	No	No	Tetraspanin	
Hayan	FBgn0030925	GH17483	No	n.s.	Yes	No	Serine protease, Innate Immune Response	
Eip55E	FBgn0000566	LD22255	No	-1.47	Yes	No	Glutathione Biosynthetic Process	
PH4alphaSG1	FBgn0051014	IP03659	No	-1.54	Yes	No	Prolyl-4-Hydroxylation (SPCG), SG-specific	
eyg	FBgn0000625	AT09010	Yes*	-1.35	No	No	Transcription Factor	Upstream and internal
sage	FBgn0037672	RE59356	Yes*	-1.96	Yes	No	Transcription Factor, SG-specific	Upstream and internal
CG13159	FBgn0033721	RE25177	Yes*	n.s.	Yes	No	Secreted protein, SG-specific	Upstream
fkh	FBgn0000659	RE03865	Yes	n.s.	No	No	Transcription Factor	Mostly upstream
sens	FBgn0002573	IP01345	Yes	n.s.	No	No	Transcription Factor	Upstream
rib	FBgn0003254	LD16058	Yes	n.s.	No	No	Transcription Factor	Upstream†
bowl	FBgn0004893	LD15350	Yes	n.s.	No	No	Transcription Factor	Upstream and internal
Xbp1	FBgn0021872	GH09250	Yes	n.s.	Yes	Yes	Transcription factor	All internal
Tudor-SN	FBgn0035121	LD20211	Yes	-1.93	Yes	Yes	RNA-induced silencing	Mostly internal
roq	FBgn0036621	LD12033	Yes	-1.32	Yes	Yes	RNA Binding Factor (Zinc Finger, Ring-type)	All internal
eIF3j	FBgn0027619	GH12681	Yes	n.s.	Yes	Yes	Translation Initiation Factor	All internal
Mvl	FBgn0011672	LD24465	Yes	-1.37	Yes	No	Metal ion transporter	All internal
mnd	FBgn0002778	LD25378	Yes	n.s.	Yes	No	Amino Acid/Polyamine (leucine) Transporter	All internal
GILT1	FBgn0038149	LD47508	Yes	n.s.	Yes	No	Disulfide oxido-reductase, lysosomal (SPCG?)	Mostly internal
Papss	FBgn0020389	LD25351	Yes	n.s.	No	No	Sulfate Adenylyl transferase, SG and epidermis (SPCG?)	All internal
CG7872	FBgn0030658	LD24870	Yes	-1.81	Yes	Yes	HSP70, ER-Resident (SPCG)	All internal
sel	FBgn0263260	GH10427	Yes	-2.87	Yes	Yes	Saposin-like, ER-Resident (SPCG)	Mostly internal
wbl	FBgn0004003	IP02648	Yes	n.s.	No	Yes	Protein Folding, ER-Resident (SPCG)	All internal
Ufc1	FBgn0034061	LD28985	Yes	n.s.	Yes	Yes	Ufmylation Transferase Activity, ER stress response (SPC	Upstream and internal
Der-1	FBgn0267972	GH08782	Yes	n.s.	No	Yes	Ubiquitin-Dependent ERAD Pathway (SPCG)	All internal
CG5021	FBgn0035944	RE32705	Yes	-1.68	Yes	Yes	Intergral Membrane, Golgi-Resident (SPCG)	All internal
CG8230	FBgn0027607	GH02536	Yes	-1.52	Yes	Yes	Dymeclin Family, Golgi-Resident (SPCG)	Upstream and internal
lqfR	FBgn0261279	GH02671	Yes	n.s.	Yes	Yes	Epsin-Like domain, Golgi-Resident (SPCG)	Upstream and internal
tbc1	FBgn0031304	AT03044	Yes	-1.23	Yes	Yes	Rab-GAP, Endosomal to Golgi trafficking, (SPCG)	Upstream and internal
Osbp	FBgn0020626	LD31802	Yes	n.s.	Yes	Yes	Oxysterol-Binding (SPCG)	All internal

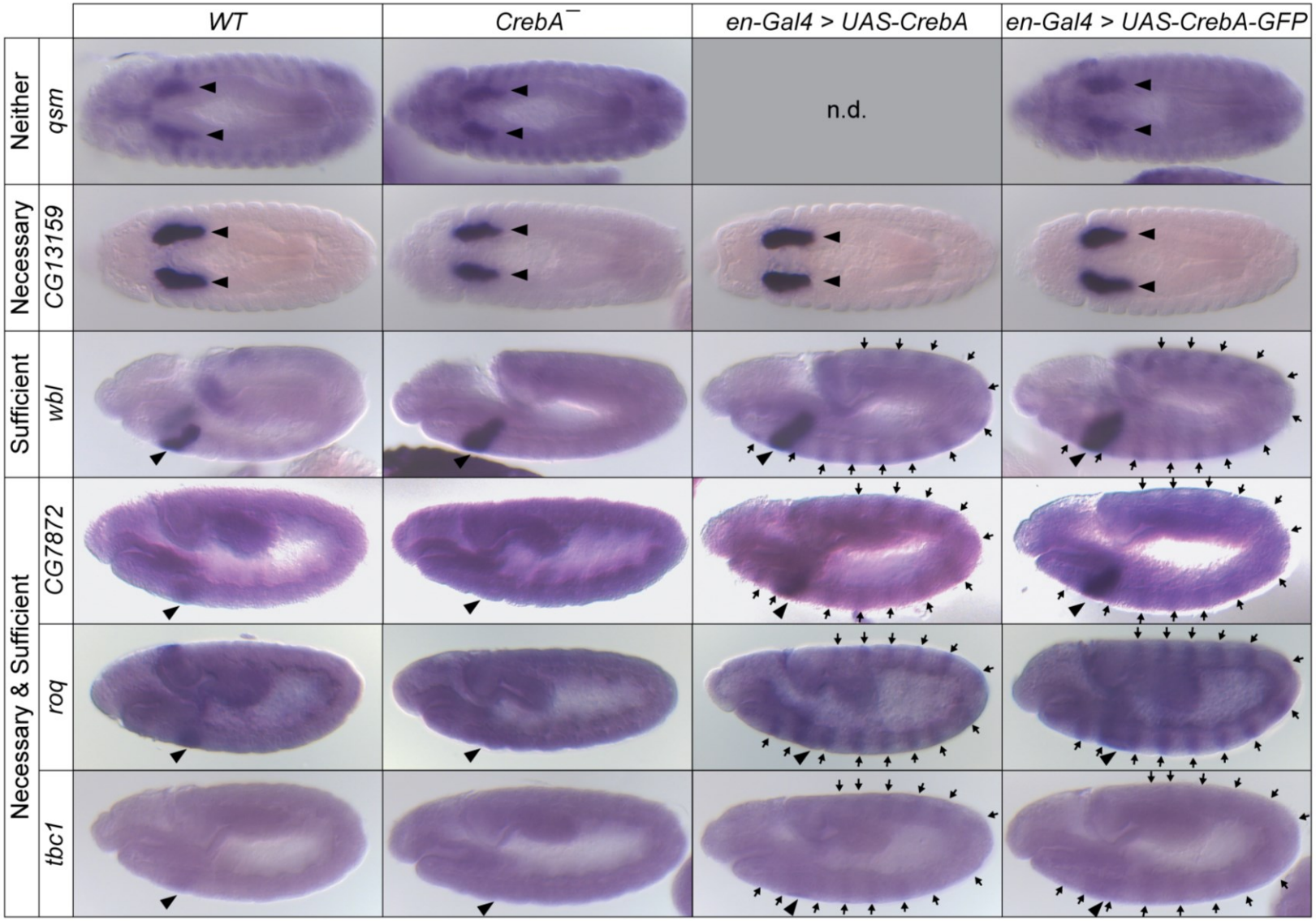
Table 1: Testing of genes for CrebA regulation.

Microarray data from Fox et al 2010; only fold changes with p-value of 0.05 or less shown. *Only with Fkh and Sage driven; †Upstream of isoform expressed in embryo

We verified SG expression for all 33 genes and found that for ten of the genes, CrebA was neither necessary nor sufficient for their SG expression (Table 1, Fig. 8). These ten genes include four for which there was no CrebA binding and six for which binding was detected in all three ChIP experiments. Notably, more than half of the genes that were completely unaffected by loss or overexpression of CrebA encode transcription factors, the class of genes most highly represented in the DAVID analysis of ChIP-Seq bound genes (Fig. 5E).

CrebA was necessary for full SG expression of nine of the 33 genes we tested but was not sufficient, based on the absence of striped expression when CrebA expression was driven in ectopic stripes. This class included two additional transcription factors genes, *Aatf* and *sage*. Notably, CrebA was both necessary and sufficient for expression of almost all known SPCGs. Four additional genes, not previously known as SPCGs, were similarly affected by loss and misexpression of CrebA: a transcription factor gene – *Xbp1*, two genes encoding proteins known to associate with RNA – *Tudor-SN* (*TSN*) and *Roquin*, and a gene encoding a translation initiation factor – *EIF3j*. For all but one of the genes for which previous microarray studies (of whole embryos) had indicated significant decreases in expression in *CrebA* mutants, decreases were also observed in the *in situ* analyses (Table 1). Importantly, the expression patterns observed with <u>all</u> of the genes affected by either CrebA loss and/or overexpression indicate that CrebA functions as a transcriptional activator. The failure to observe changes in gene expression patterns in all genes bound by CrebA when *CrebA* function is missing or when *CrebA* is mis-expressed indicates that CrebA binding does not always indicate CrebA regulation. Finding genes that are not bound by CrebA based on the ChIP-Seq analysis but whose expression nonetheless changes with CrebA loss, suggests that CrebA can also indirectly regulate SG gene expression, perhaps through other CrebA-dependent transcription factors.

Figure 8: Subset of *in situ* hybridization of genes from Table 1.Arrowhead: SG; Arrow: RNA detected in stripes of cells expressing *UAS-CrebA* and/or *UAS-CrebA-GFP* under the control of *en-Gal4*. All genes are expressed or upregulated in the SG. There is no overt change in *qsm* and *wbl* expression in *CrebA*$^{-}$ embryos; *wbl* is ectopically expressed in en stripes when CrebA is (both untagged and GFP tagged). *qsm* and *CG13159* do not have ectopic en stripe expression, *CG13159* is down regulated in the *CrebA*$^{-}$ embryos CG7872, *roq*, and *tbc1* are downregulated in *CrebA*$^{-}$ and upregulated in the en stripes when CrebA is driven in those stripes.


WT
CrebA⁻
en-Gal4 > UAS-CrebA
en-Gal4 > UAS-CrebA-GFP
Neither
qsm
n.d.
Necessary
CG13159
Sufficient
wbl
Necessary & Sufficient
CG7872
roq
tbc1

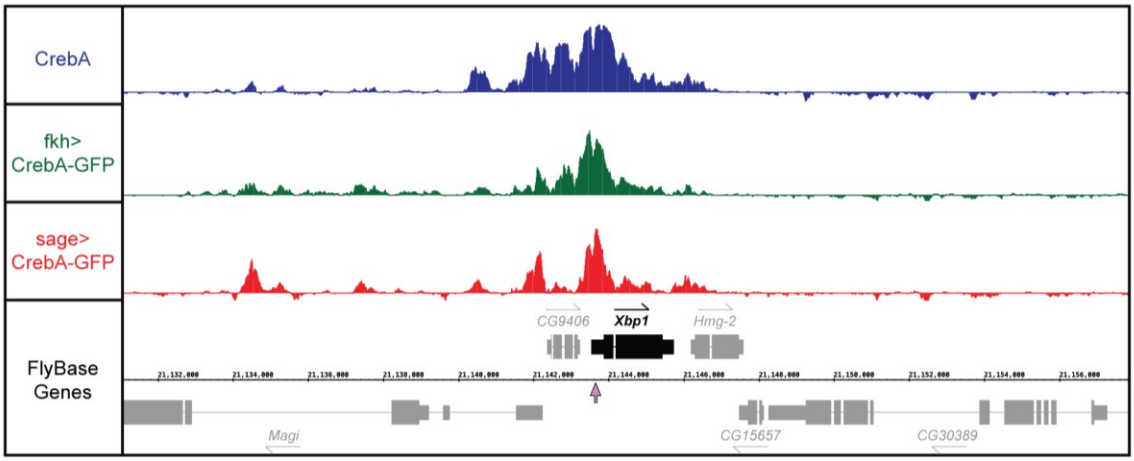
CrebA binding and CrebA regulation requires the CrebA consensus sites

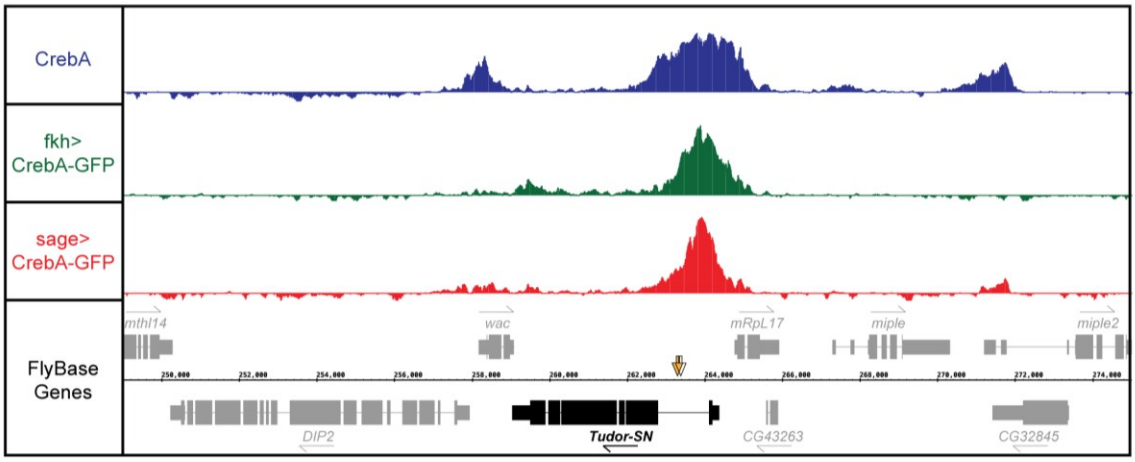
Our next experiments explored the requirements for the consensus CrebA binding motifs for regulation of two of the newly identified genes whose expression is fully responsive to changes in CrebA, *Xbp1* and *Tudor-SN* (See Table 1). Using CRISPR/Cas9-mediated recombination, we mutated the central four nucleotides of both consensus CrebA binding motifs centered in the region of the *Xbp1* gene that was pulled down in the ChIP-Seq experiments. We used a similar approach to mutate two CrebA binding motifs centered in the region of the *TSN* gene that was pulled down in the ChIP-Seq experiments, but succeeded in mutating only one of the two consensus sites. The CrebA consensus binding sites map within the 5' UTR of *Xbp1* and within the first intron of *TSN* (Fig. 9). We refer to the CrebA motif mutant alleles as $Xbp1^{CMM}$ and TSN^{CMM}. The $Xbp1^{CMM}$ and TSN^{CMM} alleles were balanced and crossed into the appropriate genetic backgrounds. X*bp1* mRNA expression in WT embryos and in $Xbp1^{CMM}$ homozygous embryos was compared. We also examined expression of *Xbp1* from both alleles in embryos expressing UAS-CrebA-GFP under the control of en-Gal4. With the WT versions of *Xbp1*, we observed *Xbp1* mRNA expression in the developing SGs of WT embryos as well as in ectopic stripes in embryos expressing CrebA-GFP under the control of en-Gal4. $Xbp1^{CMM}$ resulted in a loss of SG *Xbp1* staining and no ectopic stripes of *Xbp1* expression in embryos expressing CrebA-GFP in stripes (Fig. 10). Similar experiments are in progress with. Future studies will include using these binding site mutant lines for ChIP-qPCR with the CrebA antiserum to compare CrebA binding between the WT alleles and CMM alleles of both genes. These experiments will provide a measure of how much the presence of consensus CrebA binding motifs contributes to binding and will allow us to correlate CrebA binding and CrebA regulation.

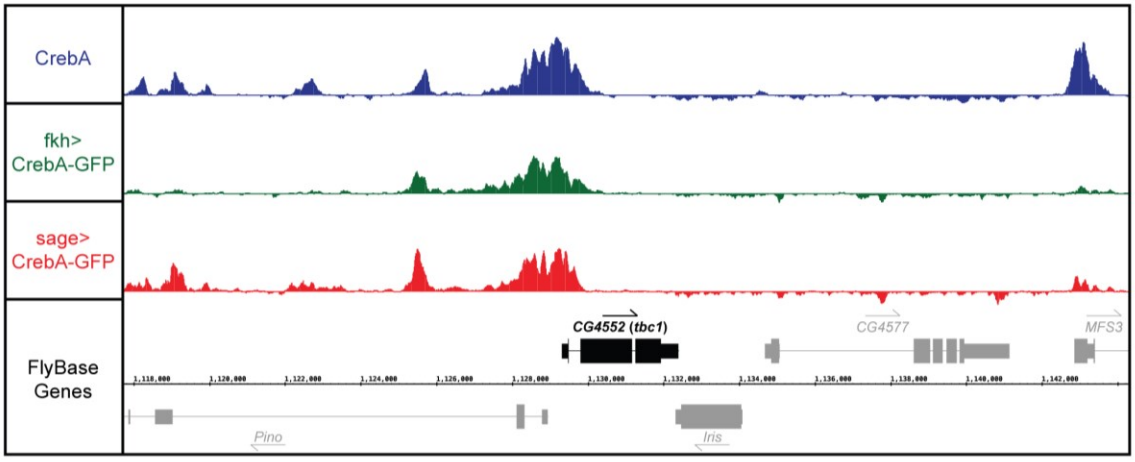
Xbp1, TSN and Tbc1 are required for SG secretory function

The role of Xbp1, TSN, and Tbc1 in SG secretion is unknown. Xbp1 has been shown to be involved in the unfolded protein response (UPR). A physiological role for this

Figure 9: ChIP-seq peaks for *Xbp1*, *TSN*, and *tbc1*. CrebA(blue) is pulldown in wild type using CrebA antisera. Arrows indicate location of CrebA binding sites. Xbp1 has two but they are 40bp apart. Filled arrows are those we successfully mutated and the empty arrowhead is the one that was not mutated.


CrebA
fkh>
CrebA-GFP
sage>
CrebA-GFP
FlyBase
Genes
CG9406
Xbp1
Hmg-2
Magi
CG15657
CG30389


CrebA
fkh>
CrebA-GFP
sage>
CrebA-GFP
FlyBase
Genes
mthl14
wac
mRpL17
miple
miple2
DIP2
Tudor-SN
CG43263
CG32845


CrebA
fkh>
CrebA-GFP
sage>
CrebA-GFP
FlyBase
Genes
CG4552 (tbc1)
CG4577
MFS3
Pino
Iris

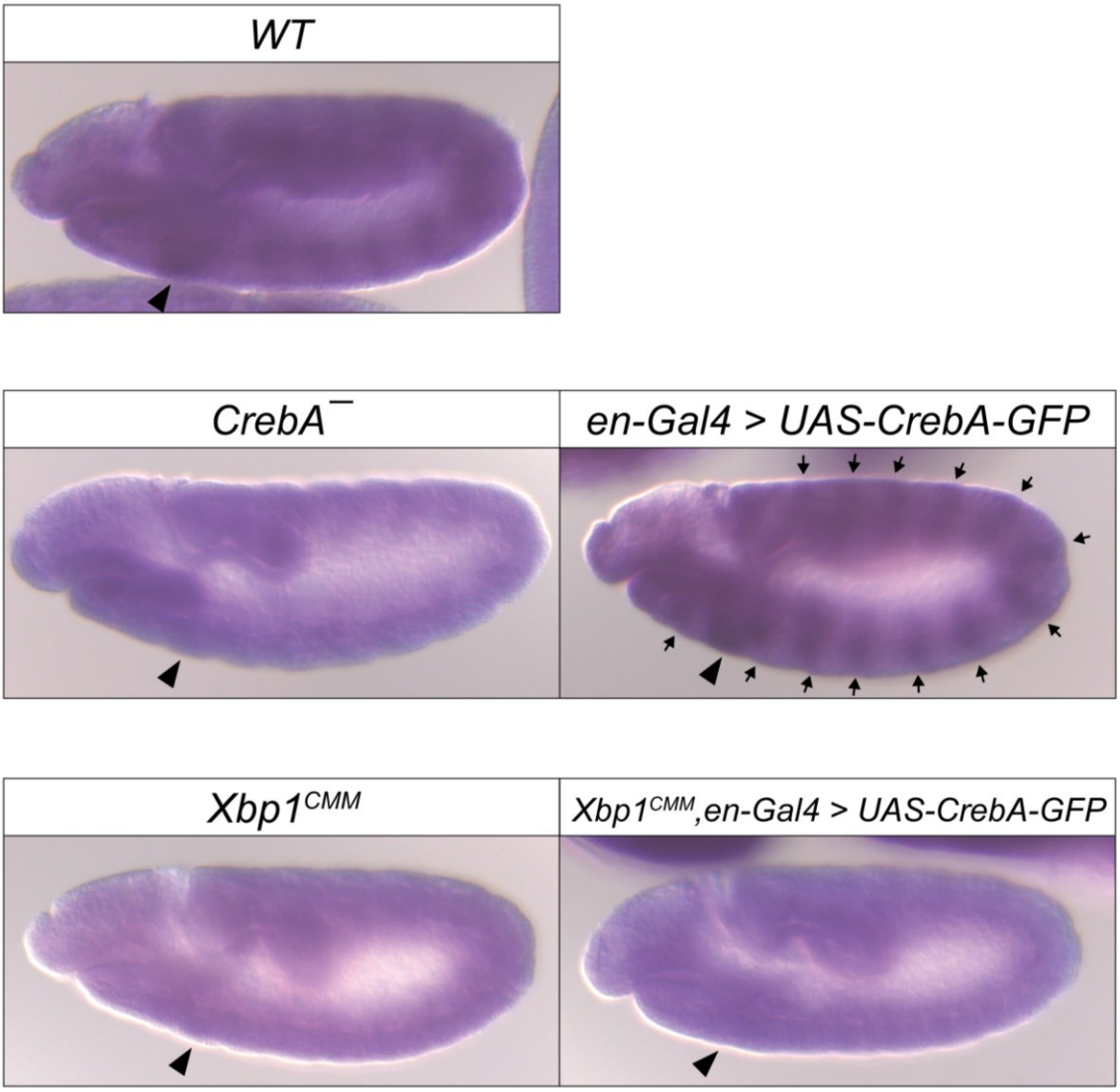


Figure 10: Xbp1 expression requires CrebA binding sites. Arrowhead: SG. Arrows: arrows indicate RNA detected in stripes of cells expressing *UAS-CrebA-GFP* under the control of *en-Gal4*.

protein is less well understood, but it stands to reason that in a dedicated secretory tissue like the SG, there would be need for upregulation of the genes found in UPR. Indeed, since proteolytic cleavage and activation of the mammalian Creb3-like proteins also occurs during UPR, many considered these proteins to be UPR-specific at least until their loss-of-function phenotypes were revealed (Murakami et al., 2009; Saito et al., 2009). TSN has been linked to cell cycle progression, where it is thought to partner with E2F-1 to drive the transition from G1 to S phase (Su et al., 2015). More recent work shows that it is the TSN degradation of miRNAs that promotes cell cycle progression (Elbarbary et al., 2017). More consistent with a potential link between TSN and CrebA function is the demonstration that, in Arabidopsis, TSN binds to and stabilizes stress-responsive mRNAs that encode secreted proteins (dit Frey et al., 2010) However, TSN has also been linked to increases in milk secretion through its proposed transcriptional activity in the nucleus (Ao et al., 2015). Thus, to learn how and where TSN functions, we generated null mutations by homologous recombination and we made antiserum that detects the endogenous protein (Fig. 11). Our immuno-staining reveals that TSN localizes to the ER (Fig. 12), a prime location to regulate secretion and/or secretory capacity. Recent characterization of the mammalian homolog of Tbc1, Tbc1d23, reveals a role for this protein in linking endocytic vesicles to the trans Golgi membrane through its binding to either of two *trans* Golgi proteins and the WAVE complex (Shin et al., 2017). Thus, Tbc-1 could also be required for normal secretory function in the SG. Thus, we generated a null allele of *tbc1* by homologous recombination (Fig. 13) and we generated antiserum to the C-terminal half of the protein (see Chapter 3 for details on the Tbc1 antiserum). Co-staining with antibodies to a *cis* Golgi marker (GM130) and a *trans* Golgi marker (Golgin-245) revealed that endogenous Tbc1 made “twin spots” with GM130 and partially co-localized with Golgin-245 (Fig. 14). We frequently observed staining where GM130 punctae would be next to Golgin-245 punctae that, in turn, would either overlap with or be next to Tbc1 punctae, suggesting that Tbc1 localizes to the trans most portion of the TGN. We also observed Tbc1 staining in vesicles near the apical surface that overlapped staining with the apical

Figure 11. **TSN Knockout generation**. A) Genomic location of *TSN*. The deficiency BSC125 indicated. Small arrows indicate direction of transcription. The region removed in the knockout is indicated. TSN KO F and R primers were used with pw 25-2 and -3, respectively, to confirm the knockout. The thick black line represents the homologous regions used in the knockout procedure. B) PCR of wild type and the TSN knockout line. Primers used are those indicated in A. C) *In situ* stains of *TSN* and *lacZ* mRNA in heterozygous and homozygous *TSN* knockout siblings. Green arrows indicate LacZ expression from the balancer chromosome; white arrow indicates the salivary gland. *TSN* mRNA is absent when there is no *lacZ* mRNA (i.e in homozygous knockout embryos).

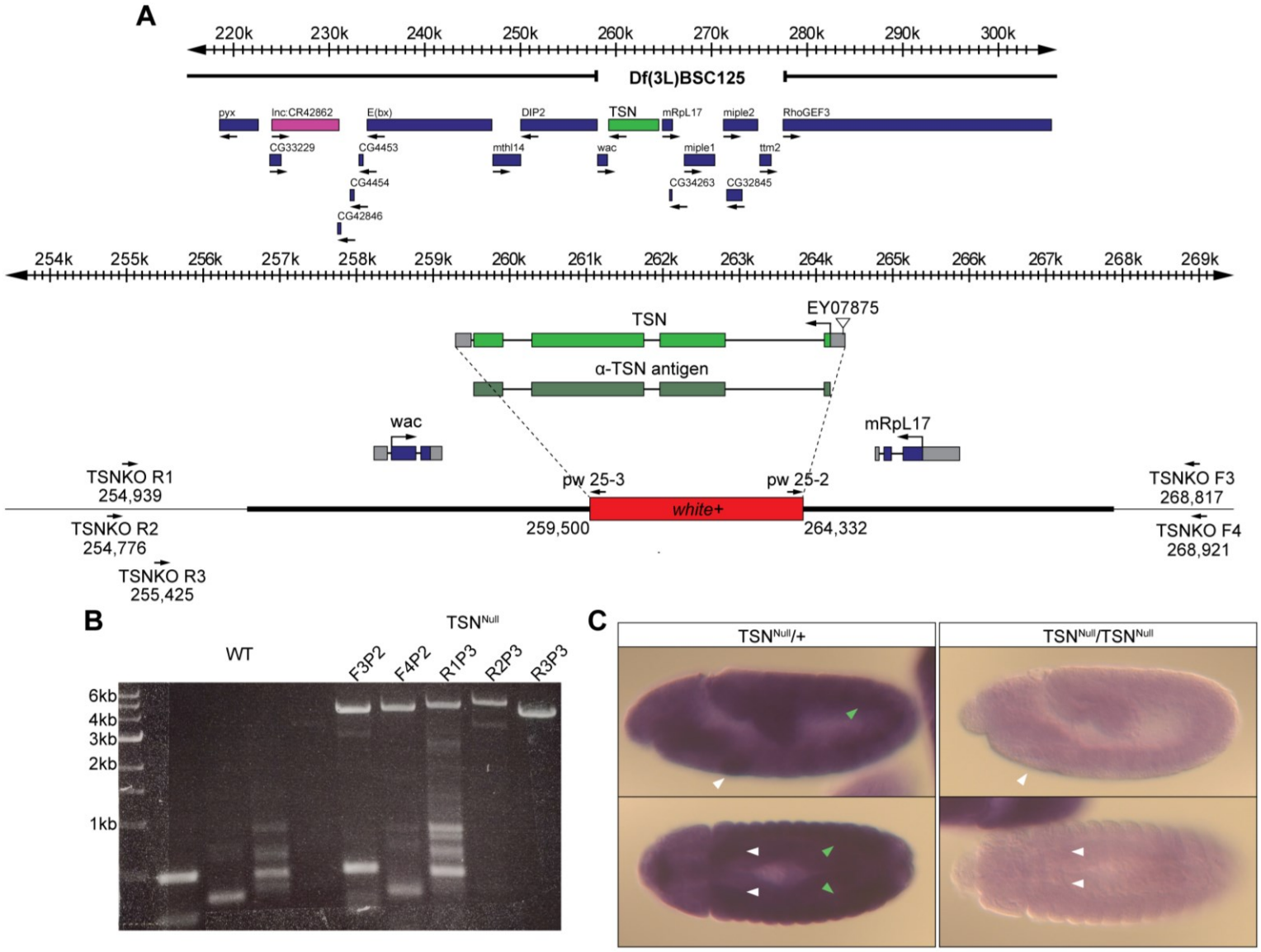

A
220k
230k
240k
250k
260k
270k
280k
290k
300k
Df(3L)BSC125
pyx
lnc:CR42862
E(bx)
DIP2
TSN
mRpL17
miple2
RhoGEF3
CG33229
CG4453
mthl14
wac
miple1
ttm2
CG4454
CG34263
CG32845
CG42846
254k
255k
256k
257k
258k
259k
260k
261k
262k
263k
264k
265k
266k
267k
268k
269k
TSN
EY07875
α-TSN antigen
wac
mRpL17
TSNKO R1
254,939
TSNKO R2
254,776
TSNKO R3
255,425
pw 25-3
pw 25-2
white+
259,500
264,332
TSNKO F3
268,817
TSNKO F4
268,921
B
WT
TSNNull
F3P2
F4P2
R1P3
R2P3
R3P3
6kb
4kb
3kb
2kb
1kb
C
TSNNull/+
TSNNull/TSNNull

Figure 12. **TSN antibody reveals that endogenous TSN and GFP-TSN localize to the ER**. A) fkh-Gal4 driver embryos stained for TSN (A; red) and the ER marker SG2 (A'; green) and merged with DAPI staining (A"; blue). B) fkh-gal4 driven UAS-GFP-TSN stained for TSN (B; red), and GFP (B'; green) and merged with DAPI staining (B"; blue). C) fkh-gal4 driven UAS-GFP-TSN stained for TSN (C; red), and SG2/ER (C'; green) and merged with DAPI staining (C"; blue).

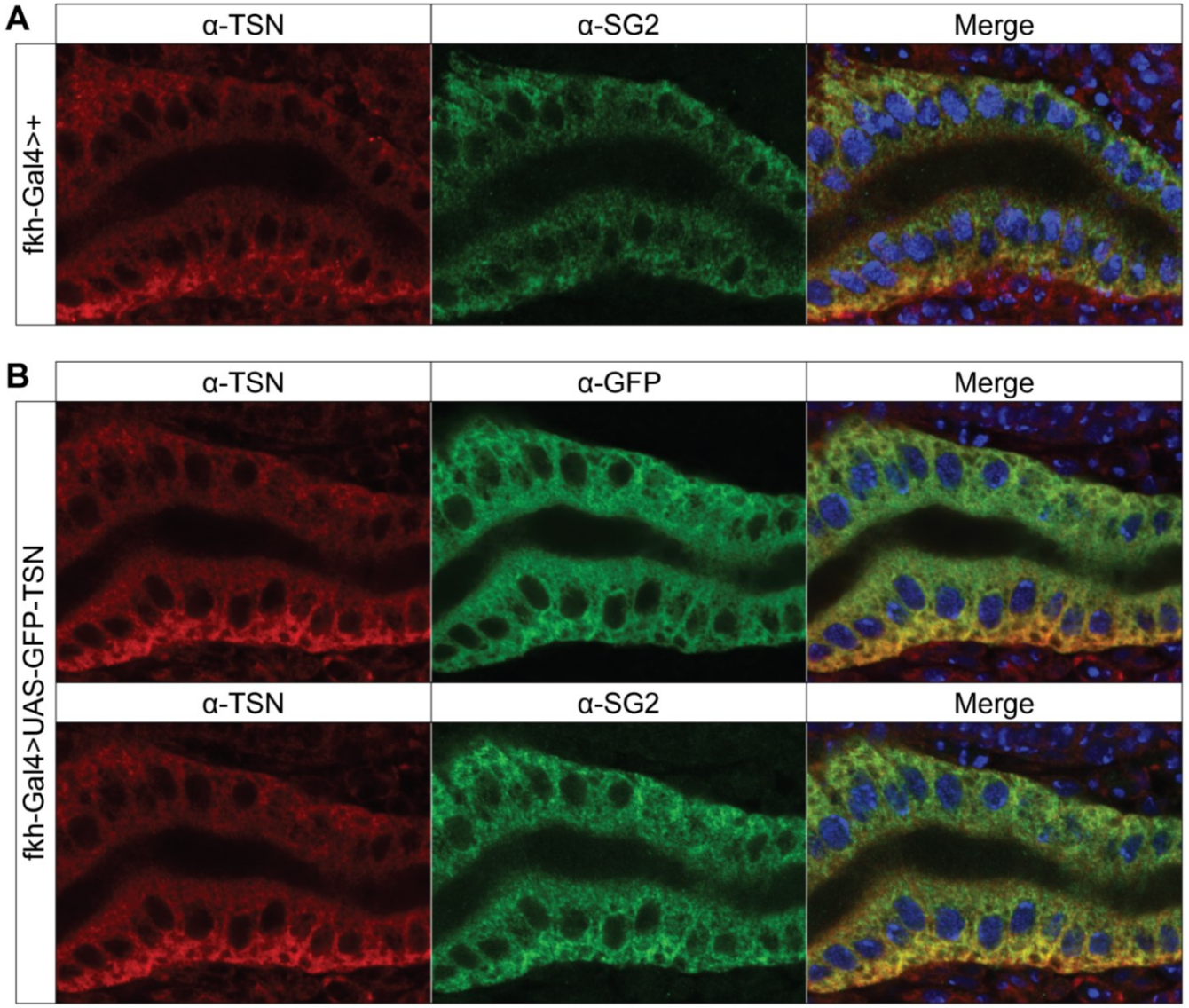

A
α-TSN
α-SG2
Merge
fkh-Gal4>+
B
α-TSN
α-GFP
Merge
fkh-Gal4>UAS-GFP-TSN
α-TSN
α-SG2
Merge

Figure 13: **Tbc1 Knockout construct**. A) Genomic location of Tbc1. The deficiency Exel6004, used in later studies, is indicated. Small arrows indicate direction of transcription. The region removed in the knockout is indicated. Tbc1KO F and R primers were used with pw 25-2 and -3, respectively, to confirm the knockout. The thick black line represents the homologous regions used in the knockout procedure. Primers for constructing the knockout homologous regions are indicated. B) PCR of two Tbc1 knockout lines, 178 and 305. 178 was used for all the experiments throughout the rest of this work. The negative control is one of the lines that is transgenic for the knockout construct, and maps to a different chromosome. Primers used are those indicated in A. Arrow is expected band for the 5' end, arrowhead is expected band for the 3' end. Asterisks indicate bands found in the negative control. C) *In situ* stains of *tbc1* and *lacZ* mRNA in heterozygous and homozygous *tbc1* knockout siblings. Green arrows indicate LacZ expression from the balancer chromosome; Black arrow indicates the salivary gland. *tbc1* mRNA is absent when there is no *lacZ* mRNA (i.e in homozygous knockout embryos).

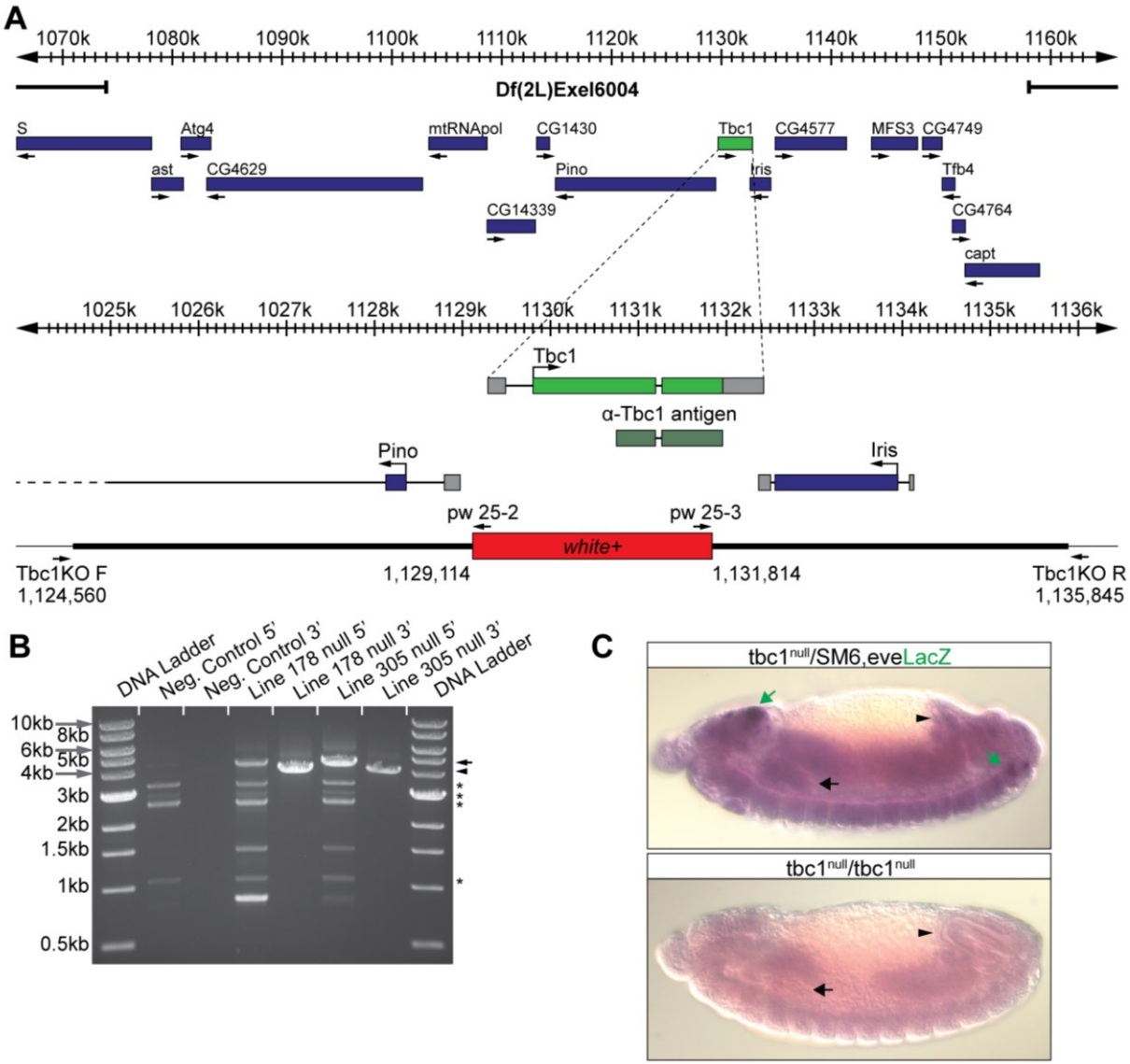
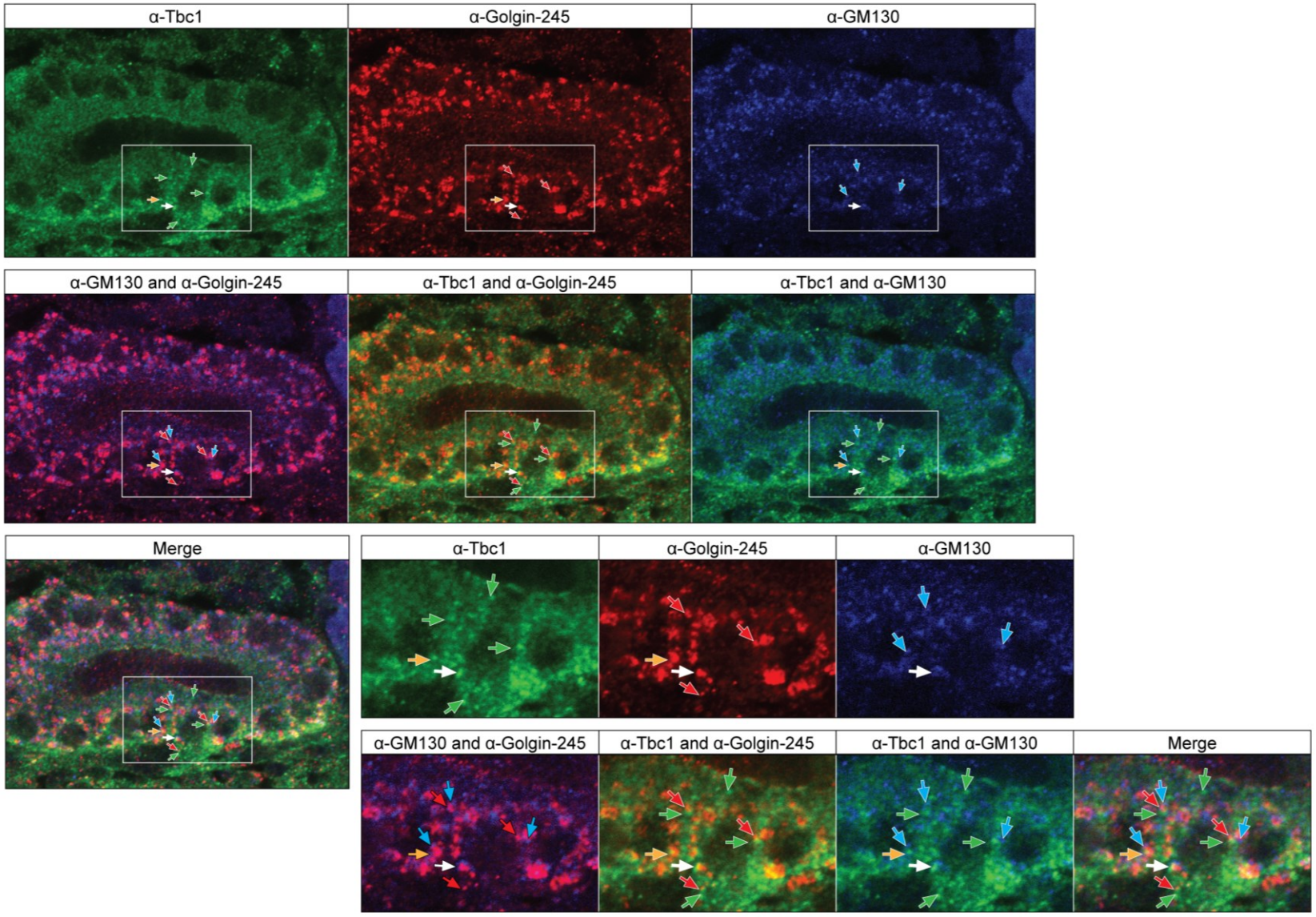

A
1070k
1080k
1090k
1100k
1110k
1120k
1130k
1140k
1150k
1160k
Df(2L)Exel6004
S
Atg4
mtRNApol
CG1430
Tbc1
CG4577
MFS3
CG4749
ast
CG4629
Pino
Iris
Tfb4
CG14339
CG4764
capt
1025k
1026k
1027k
1128k
1129k
1130k
1131k
1132k
1133k
1134k
1135k
1136k
Tbc1
α-Tbc1 antigen
Pino
Iris
pw 25-2
pw 25-3
white+
Tbc1KO F
1,124,560
1,129,114
1,131,814
Tbc1KO R
1,135,845
B
DNA Ladder
Neg. Control 5'
Neg. Control 3'
Line 178 null 5'
Line 178 null 3'
Line 305 null 5'
Line 305 null 3'
DNA Ladder
10kb
8kb
6kb
5kb
4kb
3kb
2kb
1.5kb
1kb
0.5kb
C
tbc1null/SM6,eveLacZ
tbc1null/tbc1null

Figure 14: Localization of Tbc1 to the trans most part of the TGN. Tbc1 (Green), Golgin-245 (Red) and GM130 (Blue) staining of a WT SG. Boxes mark the higher resolution images. Green arrows: Tbc1; Red arrows: Golgin-245; Blue arrows: GM130; Orange arrows: Overlapped Tbc1 and Golgin-245 punctae; white arrow: overlap of Tbc1, Golgin-245, and GM130.


α-Tbc1
α-Golgin-245
α-GM130
α-GM130 and α-Golgin-245
α-Tbc1 and α-Golgin-245
α-Tbc1 and α-GM130
Merge
α-Tbc1
α-Golgin-245
α-GM130
α-GM130 and α-Golgin-245
α-Tbc1 and α-Golgin-245
α-Tbc1 and α-GM130
Merge

recycling endosome marker Rab11 (van IJzendoorn, 2006).

To determine if *Xbp1*, *TSN*, and *tbc1* effect SG secretion, we stained homozygous mutant embryos with a polyclonal antibody generated to the nuclear En protein, EnP1,but that also has off-target staining in the SG lumen and SG secretory vesicles (Myat and Andrew, 2002). By stage 16 of embryogenesis, wild-type embryos have a full lumen with little to no vesicular staining in the cells (Fig 15A; Fig 16). CrebA null mutant embryos have very thin, sometimes crooked, lumenal SG staining domains with no detectable cellular staining (Fig 15B; Fig 16). Embryos homozygous for the $Xbp1^{CMM}$ allele, have SGs with very diffuse cellular and lumenal staining (Fig 15C; Fig 16). We also often observe samples where the luminal content appears to have pulled away from the apical membranes. We compared SGs from homozygous $tbc1^{Null}$ embryos from heterozygous mothers to SGs from homozygous $tbc1^{Null}$ mutants whose mothers were also homozygous for $tbc1^{Null}$. $tbc1^{Null}$ maternal/zygotic and $tbc1^{Null}$ zygotic embryos showed an increase in light and dark cellular staining (Fig 15D,E; Fig 16), with the maternal/zygotic mutants having a slightly stronger, albeit not significant, phenotype. The presence of cellular staining suggests that the EnP1-positive lumenal contents are not being trafficked to the cell surface and secreted as efficiently as in wild type SGs.

The SG secretory phenotypes associated with complete loss of *TSN* as well as with the TSN^{CMM} mutants are currently being examined.

Xbp1, TSN and Tbc1 also have a role in cuticle secretion

It was previously shown that regulators and effectors of secretion lead to deformation or loss of cuticle (Fox and Andrew, 2015). For example, both *CrebA* null mutants and embryos homozygous for mutations in genes encoding the protein components of the secretory machinery have defects in the cuticle secreted by epithelial cells (Abrams and Andrew, 2005). The phenotypes include “ghost-like” cuticles with little to no denticles and hairs, very light, underdeveloped mouthparts, and a barely detectable filzkörper (the elongated air filter found at

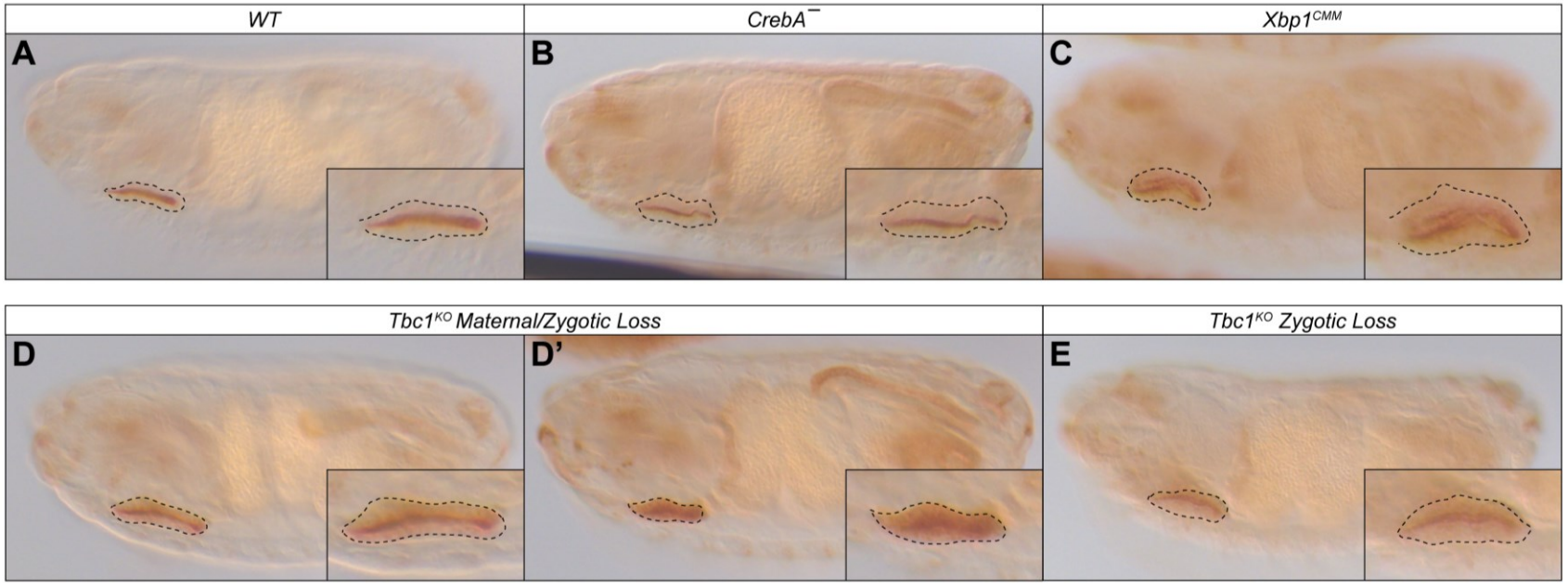


Figure 15. Staining of SG lumen. A-H) Represented stainings using an en polyclonal which also stains the SG lumen. A) Wild type, B) *CrebA* null, C) Xbp1^{CMM}, D) Tbc1^{Null} Maternal/Zygotic loss, E) Tbc1^{Null} Zygotic loss. Insets are magnifications of SG. SGs outlined with dashed lines.

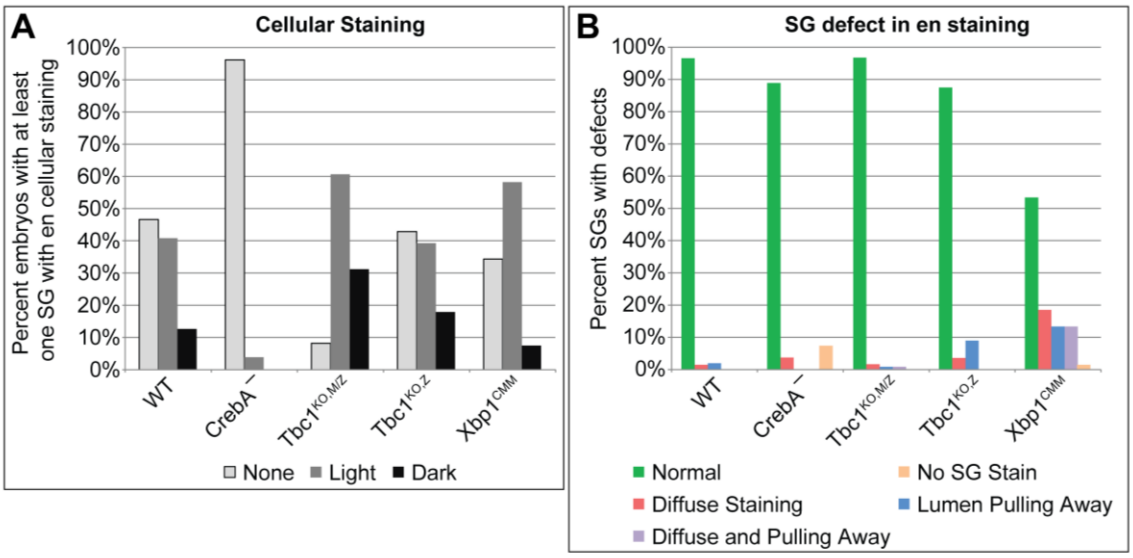
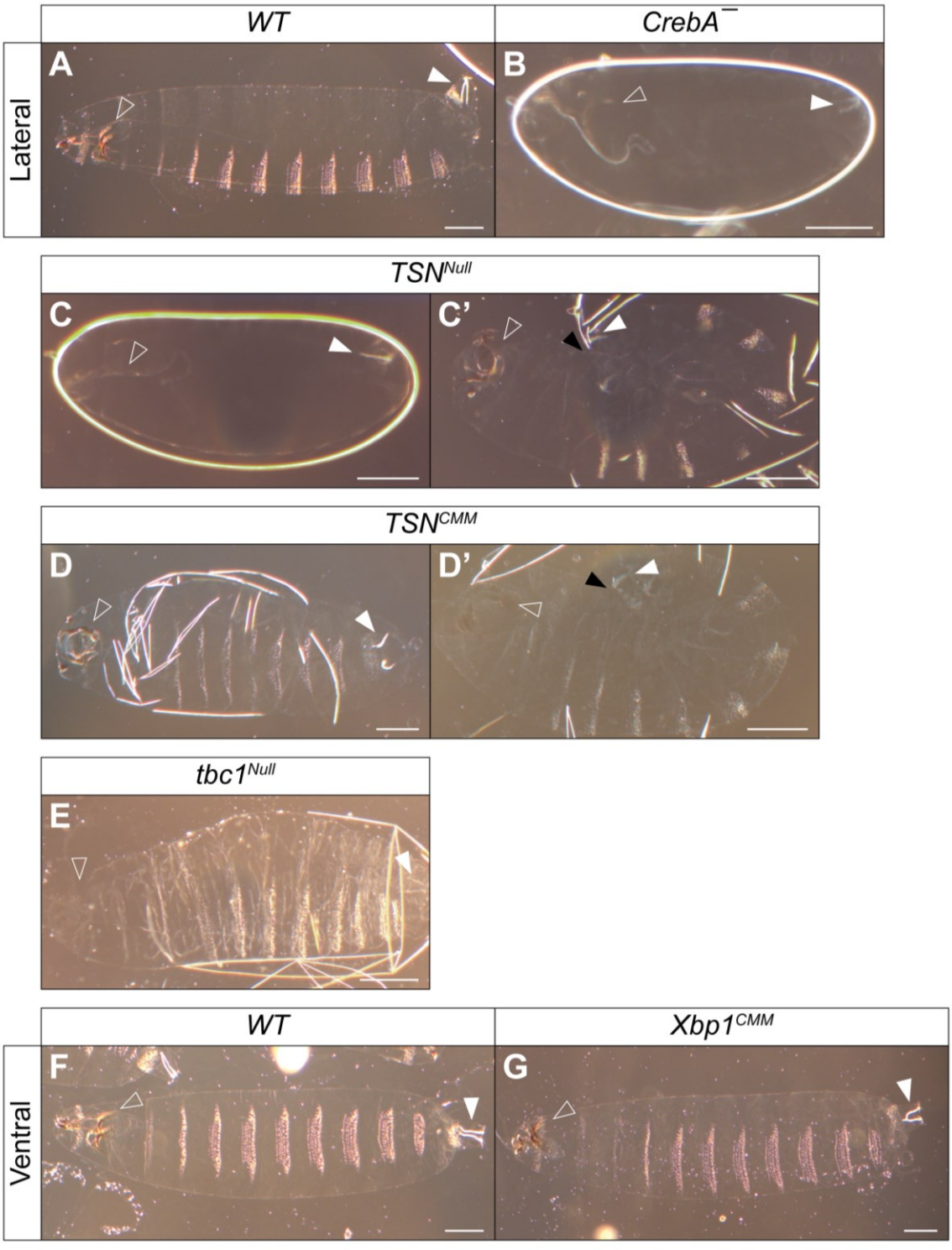


Figure 16: Quantification of SG defects. A) Quantification of the number of embryos with at least one SG with cellular staining and whether the cellular staining was light or dark. B) Quantification of SGs based on if they had diffuse staining, the lumen had pulled away, and/or no staining.

the end of the trachea, which is filled with fine cuticular threads). To determine if *Xbp1*, *TSN*, and *tbc1* effect secretion, cuticle preparations were made with the CrebA binding site mutants of *Xbp1* and *TSN* as well as knock-out mutations of *TSN* and *tbc1*. As shown previously, CrebA mutants have very little cuticle (a "ghost-like" appearance) and no obvious mouthparts or denticles (Fig. 17B). Both *TSN* knock out larvae and TSN^{CMM} larvae have ghost-like cuticles quite similar to those observed in CrebA mutants (Fig. 17C,D). We also observe a high frequency of U-shaped embryos, where the posterior end is folded back over the more anterior part of the larva (Fig. 17C,D). *Xbp1* binding site mutants and *tbc1* knockout mutants have more subtle defects with mouthparts that are shorter and heads that are thinner than in wild type (Fig. 17E-G). Both sets of homozygous mutants had a significant increase in the observed cuticle defects. As the mouthparts form from secreted cuticle, the defects suggest that both *Xbp1* and *tbc1* function in the epidermal secretion of larval cuticle, although their contributions to the formation of the cuticle is not as critical as the contributions from CrebA and TSN.

Figure 17. Cuticle preparations of candidates show a range of secretion defects. A-G) Representitive cuticle preps. A) Wild type-Lateral View, B) *CrebA* null C) TSN^{Null} D) TSN^{CMM}, E) Wild Type-Ventral View, F) $Xbp1^{CMM}$, and G) $Tbc1^{Null}$ Maternal/Zygotic loss. Arrowhead outline: mouthparts. White arrowhead: U: shaped bend. Scale bar: 100μm.


WT
CrebA⁻
Lateral
A
B
TSN^Null
C
C'
TSN^CMM
D
D'
tbc1^Null
E
WT
Xbp1^CMM
Ventral
F
G

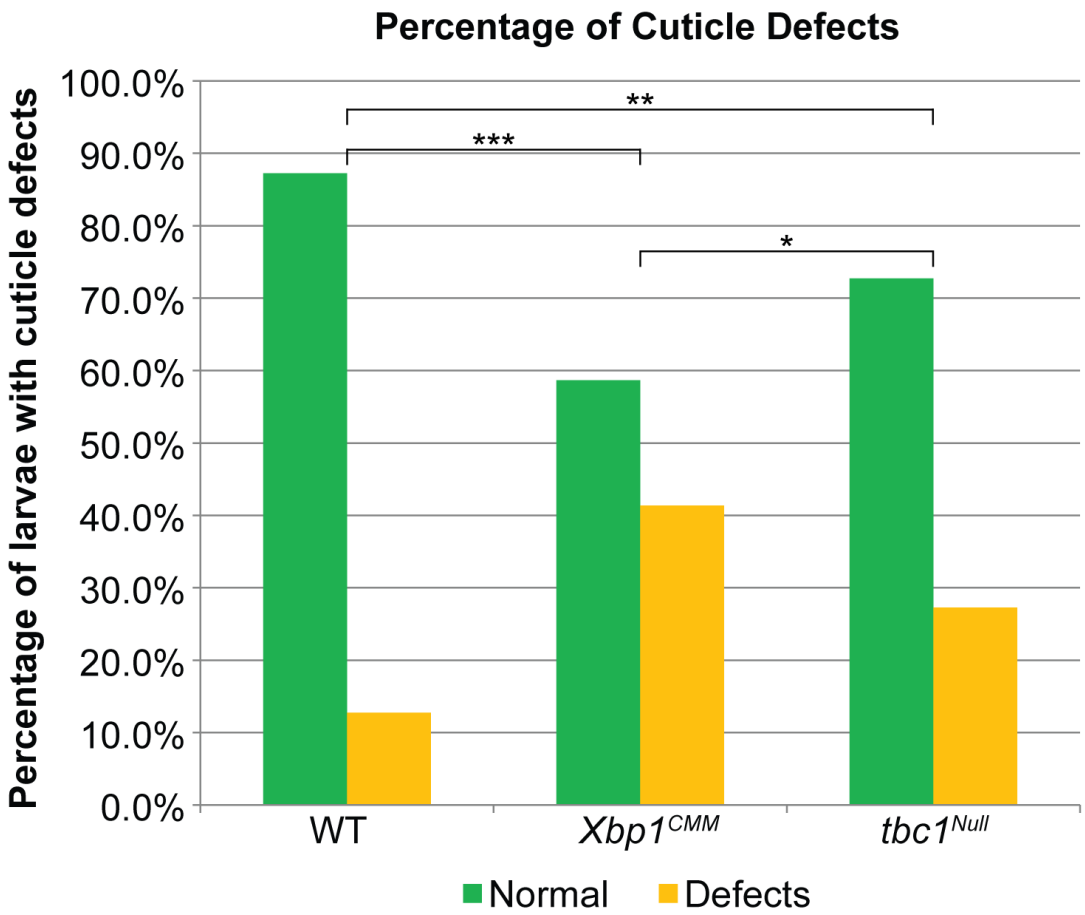


Figure 18: Quantification of Cuticle Defects of *Xbp1*^{CMM} and *tbc1*^{Null}. *: p-value=0.02. **: p-value=0.002. ***: p-value=$1.4*10^{-24}$.

Discussion

Using a fully functional CrebA tagged with GFP (Fig. 3), we performed ChIP-seq to reveal the SG-specific binding of CrebA. To this end, we utilized two drivers that only had the SG in common and studied the binding peaks that overlapped both datasets. In conjunction, we performed ChIP-seq on wild-type embryos using CrebA antisera (Fig. 2). By comparing all three datasets, we determined that SG binding represents a large portion of the total CrebA binding and that CrebA exhibits tissue-specific binding (Fig. 5,6). Analysis of the motifs present in the bound DNA revealed enrichment for the consensus binding sites for the mammalian Creb3-like and Drosophila CrebA that had been determined from *in vitro* studies. The mammalian Xbp1 consensus-binding site, which is similar to that of the Creb3-like/CrebA proteins, is also enriched. CrebA consensus binding motifs are enriched around the binding peak summits from the ChIP-Seq analysis, suggesting that the *in vitro* CrebA binding motifs also function *in vivo*.

Potential target genes associated with the SG-specific *in vivo* CrebA binding sites were identified and a GO term analysis was performed to reveal the classes of genes associated with CrebA binding. This analysis revealed that TFs and other nucleic acid binding proteins were the most enriched, regardless of which dataset intersection was used (Fig. 5,6). These target gene classes are inconsistent with previous expression studies that have shown that CrebA primarily regulates SPCGs and secretory cargo (Fox et al., 2010), bringing up the question of whether binding is always linked to CrebA regulation. CrebA binds almost half of the genes expressed in the SG based on BDGP *in situ* analysis, suggesting that CrebA binding may be linked to open chromatin. Consistent with this idea, immuno-staining of SG polytene chromosomes with CrebA antiserum reveals that CrebA binds to low DAPI intensity regions (Fig. 7E). Furthermore, a CrebA consensus site is associated with between 74.5 and 94.5% of the binding events in the SG (depending on which of the three overlapping datasets is analyzed). These observations suggest that CrebA binding is a consequence of chromatin accessibility combined with the presence of

consensus CrebA binding sites, and that CrebA binding may not always indicate CrebA regulation.

Our *in situ* analysis of 33 newly tested SG genes – 25 bound by CrebA in the SG and eight not bound – supports the idea that CrebA binding and CrebA regulation can be unlinked (Table 1, Fig. 8). A breakdown of the 25 genes bound by CrebA in the SG shows some interesting patterns. Of the seven SG TFs bound by CrebA (Table 1, green shading), only two show reduced expression in the absence of CrebA (*sage* and *Xbp1*), and only expression of *Xbp1* was induced in the ectopic stripes of CrebA expression. Of the ten SG SPCGs bound by CrebA (Table 1, dark blue shading), most showed reduced expression in the absence of CrebA (8/10) and expression of all ten was induced in the ectopic stripes of CrebA expression. The ten genes include *tbc1* – a gene whose mammalian orthologue was linked to endosomal to TGN trafficking while our studies were ongoing. Of the eight SG-expressed genes that were not bound by CrebA in the SG, four showed reduced expression in *CrebA* null SGs but none were activated by ectopic *CrebA*. These findings suggest that (1) CrebA binding is essential for CrebA-dependent ectopic expression, (2) CrebA upregulation of a subset of target genes is indirect and possibly through either Sage or Xbp1 (the two SG transcription factors whose SG expression was reduced with *CrebA* loss), (3) that a significant subset of SG genes bound by CrebA are not regulated by CrebA (6/25 CrebA bound SG expressed genes were completely unaffected by CrebA loss or overexpression). This analysis also fully supports a direct role for CrebA in regulation of SPCGs. Expression of all of the *bona fide* SPCGs bound by CrebA (10/10) were ectopically induced in the stripes of ectopic CrebA expression. This finding suggests that some of the other genes for which CrebA is sufficient for their expression, including *Xbp1*, *TSN*, *roq*, and *elF3j*, may play major roles in secretion. We demonstrate this is the case for Xbp1 and TSN.

In this study, we also test the importance of the CrebA consensus binding sites for the regulation of two new direct targets of CrebA, both of which are shown here to encode regulators of secretory capacity, *Xbp1* and *TSN*. With *Xbp1*, mutations in the two CrebA consensus-binding

sites led to loss of CrebA regulation of *Xbp1.* Both CrebA-dependent endogenous SG expression was lost as well as the ectopic stripes of *Xbp1* observed with the ectopic striped expression of CrebA. We are currently testing if these same mutations result in loss of CrebA binding, which we predict will be the case.

The connection of CrebA to the classic UPR regulator *Xbp1* is extremely interesting as the physiological role of the UPR is still being elucidated. Many studies of the UPR have been by using exogenous application of chemicals such as thapsigargin (Shinjo et al., 2013). While invaluable in understanding the mechanisms of the UPR, it is unlikely that the UPR evolved in response to chemical perturbation of ER protein folding. Rather, we propose that the UPR exists to ensure efficient processing and exocytosis of secreted cargo. *Xbp1* activation during UPR occurs by alternative splicing by the canonical UPR regulator IRE-1, resulting in a frameshift that produces functional Xbp1 (Frakes and Dillin, 2017). Xbp1 is a transcriptional regulator and is known to regulate multiple UPR processes including upregulation of chaperone genes. In the role as a regulator of UPR, Xbp1 has been implicated in the immune response in *C. elegans* where mutants have decreased survival when exposed to a bacterial pathogen and abnormal ER morphology (Frakes and Dillin, 2017). A subset of patients with inflammatory bowel disease carry Xbp1 hypomorphic alleles. Loss of Xbp1 during mammalian embryonic development is lethal. Rescue in the liver rescues embryonic lethality, but the mice die a few days after birth due to secretion defects in both the SG and pancreas. Even heterozygous Xbp1 null mice are not perfectly healthy, developing diabetes and other diabetic symptoms such as insulin insensitivity. Expression of Xbp1 in neurons can lengthen the lifespan of *C. elegans*. These studies support a model in which the high secretory load in these secretory tissues, induces a form of ER stress, which then upregulates Xbp1 and related UPR genes. Considering that Xbp1 has a similar binding site motif as CrebA, Xbp1 may work with CrebA to regulate a subset of SPCGs, potentially function either as a co-activator or as a downstream TF to activate genes not bound by CrebA.

TSN has been implicated in many mRNA interactions, including RISC-mediated silencing, regulation of non-coding RNAs, and splicing (Gutierrez-Beltran et al., 2016). TSN has also been implicated in transcriptional regulation, acting as a co-activator linking TFs with the basal transcription machinery such as c-Myb. In this role as a TF, TSN has been linked to cell proliferation, differentiation, and death. Increased TSN expression is an indicator of a poor prognosis in different cancers. Indeed, TSN is even considered a good marker for breast cancer metastasis and it has been suggested that TSN could become a diagnostic or prognostic marker. With regards to a role for TSN in secretion, there are two studies in Arabidopsis and one in a bovine mammary cell line that show TSN increases secretion through two different mechanisms. In Arabidopsis, TSN acts through stabilization of mRNAs encoding secreted proteins (dit Frey et al., 2010), whereas TSN transcriptional activity has been implicated in boosting bovine milk secretion (Ao et al., 2015). Our work shows that in the embryonic SG, TSN localizes to the ER, and not the nucleus (Fig. 12), suggesting that in this tissue, TSN likely acts through mRNA regulation rather than transcription. We also show that TSN is required for efficient secretion in both the SG and in the epidermal cells that secrete the larval cuticle (Fig. 17) and that the CrebA binding sites are required for the secretory function of TSN (Fig 17). Thus, TSN functions at or near the ER to boost secretory capacity in a CrebA-dependent manner. Drosophila will be the ideal system for parsing out the underlying mechanism for this activity.

Tbc1 recently was found to function as a bridging factor linking endosomally derived vesicles to the TGN (Shin et al., 2017). As such, it was not surprising to find that CrebA is both necessary and sufficient for *tbc1* expression, since similar regulation patterns have been discovered for all other SPCGs that have been examined, with the exception of a subset that are only expressed in the SG and nowhere else, such as *PH4alphaSG1*. It was also not surprising to find that loss of Tbc1 is associated with mild secretion defects (Fig. 15, 16, 17). This is discussed more in depth in Chapter 2, where we show that the secretion role of Tbc1 is conserved.

Altogether, all three of the genes bound by CrebA and examined here have defects consistent with each playing a role in secretion. This observation suggests that most, if not all, of the genes regulated by CrebA function to increase secretory capacity. Thus, CrebA boosts secretory capacity at multiple levels, both through direct regulation of effectors such as *tbc1* and other SPCGs (Abrams and Andrew, 2005; Fox et al., 2010), and through direct upregulation of a transcription factors (Xbp1) and potential regulators of mRNA stability (TSN).

Most reported studies support a role for the entire CrebA/Creb3L family in boosting secretory capacity. Previous studies in Drosophila have linked CrebA to secretory capacity (Abrams and Andrew, 2005; Fox et al., 2010; Fox and Andrew, 2015). This study demonstrates that CrebA directly regulates most SPCGs as well as regulators of secretion. Others have shown that during dendritic arborization, CrebA upregulates the COPII machinery and that loss of CrebA or COPII components leads to a drastic decrease in neuronal arborization (Iyer et al., 2013). A recent study of the innate immune response in Drosophila shows that loss of CrebA leads to higher mortality and up-regulation of Xbp1 activation (Troha et al., 2018). Importantly, this same study used RNA-seq to reveal that CrebA upregulates SPCGs as well as a subset of secreted antimicrobial peptides (AMPs) in the fat body during infection.

In mice, the closest CrebA homologs Creb3L1 and Creb3L2 have been shown to affect osteogenesis and chondrogenesis, respectively (Saito, 2014). Both Creb3L1 and Creb3L2 have been shown to regulate expression of COPII components (Keller et al., 2017; Saito et al., 2009). Recent work reveals that a point mutation in Creb3L1 leads to familial osteogenesis imperfecta, also known as brittle bone disease (Keller et al., 2017). This specific mutation results in secretion defects due, in part, to low expression of the COPII component Sec24d. The more specific nature of loss-of-function phenotypes of individual Creb3L genes likely reflects the redundant functions of this gene family; loss of more than one of the Creb3L genes in mammals is likely to cause more profound and widespread defects such as what is observed with loss of the only CrebA gene in flies.

Materials and Methods

Fly Strains

Oregon R was used the wild-type strain. *w; UAS-CrebA* was provided by S. Smolik (Rose et al, 1997), *w; fkh-Gal4*, and *w; Sage-Gal4/CyO, ftz-lacZ*, and *Sp/ CyO, ftz-lacZ; fkh^{6} e^{3}/TM6B, Ubx-lacZ* were developed previously in the Andrew lab (Henderson and Andrew, 2000; Loganathan et al, 2016; Wiegel et al 1989). Protein null *CrebA^{wR23}* was used as the CrebA mutant throughout the text (Andrew et al, 1997). All crosses and collections were performed at room temperature (25 ºC), unless noted otherwise. The *w; en-Gal4* and *w; Sp/CyO, ftz-lacZ*; *prd-Gal4/TM6, Ubx-lacZ* lines are from Wiess et al, 2001 and Bloomington Stock Center (Indianapolis, IN; BDSC_1947), respectively. Also from the Bloomington Stock Center (Indianapolis, IN) are the lines used for balancing transgenic and other mutated lines: *w; Sco/CyO, ftz-lac*Z and *w;; TM3/TM6B, Ubx-LacZ* and *Sp/ CyO, ftz-lacZ; Dr/TM6B, Ubx-lacZ*.

Cloning

UAS-CrebA-GFP was cloned using the Gateway Cloning system. The CrebA ORF was subcloned from cDNA RE64328 into the pEnterD plasmid using TOPO cloning (Invitrogen). Site directed mutagenesis using the QuikChange kit (Aligent) was used to fix a frame shift mutation in the ORF by addition of C into position 363 (primer sequences are in Supplemental Table 1). The corrected *CrebA* ORF was then swapped into the pTWG (Gateway Cloning System; Carnegie Institution) plasmid using LR Recombination (Invitrogen); this construct adds an in-frame GFP ORF to the C-terminus of the CrebA ORF. Rainbow Transgenics, LLC injected this construct into *w^{1118}* embryos to create the transgenic insertions, which were mapped to specific chromosomes using a combination of chromosome-specific balancer lines. See fly strains.

To generate the homologous regions (HR) for CRISPR/Cas9 mutagenesis, TOPO cloning (ThermoFisher Scientific) was used to insert the genomic region into a plasmid. Site directed

mutagenesis of the CrebA binding sites was performed with the QuikChange kit (Aligent). Oregon R was the source of the gRNA to generate HR region. Primers are listed in Supplemental Table 1.

Design of gRNAs for Xbp1 and TSN mutagenesis was done with http://tools.flycrispr.molbio.wisc.edu/targetFinder/. The corresponding sequences were synthesized as DNA oligos with a 5' phosphate. Oligos were annealed at 95°C for 5 min and allowed to come to room temperature for 10 min (modified from http://tools.flycrispr.molbio.wisc.edu/; Gratz et al., 2014). The annealed oligos were ligated into linearized and dephosphorylated pU6-2-BbsI-gRNA (DGRC; Bloomington, IN; Stock #: 1363) as described in Gratz et al 2014 (http://flycrispr.molbio.wisc.edu/protocols). The entire ligation reaction was transformed into 50 µL of chemically competent DH5α *E. coli*. Injection of the gRNA, along with the homologous DNA containing the mutated CrebA consensus binding sites, into fly embryos was performed by Rainbow Transgenics, LLC (Camarillo, CA). TSN gRNAs and DNAs were injected into y^1,sc^1,v^1; *attP40{nos-Cas9}*; +/+ embryos and Xbp1 gRNAs and DNAs were injected into y^1,sc^1,v^1; *+/+; attP2{nos-Cas9}* embryos. Two lines each were established from each independent F1. Genomic DNA from these stocks was isolated and the region of interest was PCR amplified and sequenced to identify lines with mutated CrebA binding sites.

The homologous recombination knockouts of *TSN* and *tbc1* were performed as described previously (Gong and Golic, 2004). Primers used for the HR amplification and cloning as well as those used for determining if the knockout occurred at the right locus are in Table X. The HR were cloned into the pW25 vector and injected into w^{1118} embryos by Rainbow Transgenics, LLC (Camarillo, CA). The resulting transgenic insertions were mapped to specific chromosomes by tracking the $white^+$ eye color marker. An insertion line mapping to a chromosome that does not contain the endogenous gene was then crossed to a line expressing the appropriate enzymes for excision and linearization of the transgenic construct, as described in Gong and Golic (2003).

Potential knockout lines were selected for *white*$^{+}$ insertions on the chromosome containing the endogenous gene. See Fly Strains for details on the fly lines used.

Antibody Staining

Embryo fixation and immunohistochemistry were performed as previously described (Reuter and Scott, 1990) . Normal goat serum (NGS) was used except for the stains that included the donkey secondaries, where normal donkey serum (NDS) was used. Polytene chromosome preparations were performed as described previously (Fox et al., 2013). Primary and secondary antibodies, concentrations used, and sources are noted in Supplemental Table 2.

In situ hybridization

In situ hybridization was performed as previously described (Lehmann and Tautz, 1994). Antisense RNA probes for the *in situ* analysis were generated from the cDNA clones indicated in Table 1. The *Spase12* RNA probe was generated by E. Abrams (Abrams and Andrew, 2005). *LacZ* probes were also included to mark the presence of balancer chromosomes to distinguish *CrebA* homozygous embryos or embryos ectopically expressing *CrebA* from their balancer containing siblings.

Cuticle Preparation

Cuticles were prepared as described previously (Fox et al., 2010). Briefly, embryos were collected overnight on apple juice caps at 25ºC, dechorionated and aged at 18ºC for ~24 hrs. The larvae were transferred to a slide with 1:1 Hoyers:Lactic acid solution. After coverslips were placed on larvae, the slides were weighted down and put at 65ºC for 3-4 hrs.

Microscopy

All DIC and dark field microscopy was performed on a Zeiss Axiophot 2 with Janoptik ProgResC14 Plus optical imaging system. Confocal images were taken with Zeiss AxioObserver with LSM700 confocal module (JHMI Microscopy Facility, Baltimore; S10 OD016374). SIM was performed on a Nikon SMZ800.

Chromatin Preparation

Embryo collection and chromatin preparation was performed as described previously (Loganathan et al., 2016). Briefly, embryos were collected overnight on apple juice caps at 25ºC and fixed in 1.8% formaldehyde for 15 min in a 1:3 solution with heptane. All liquid was removed and the fixed embryos were then stored at -80ºC until chromatin preparation. Embryos were lysed using a dounce in a 1.8% formaldehyde solution. Chromatin was fragmented by sonication and was separated from cellular debris through serial centrifugation. Once isolated, chromatin was stored at -80ºC. The immunoprecipitation, sequencing and analysis was performed as in Loganathan et al., 2016.

Acknowledgements

We thank Melina Papadakis, Juan Guzman and Tony Zou for sequencing the TSN lines to identify the CrebA binding site mutants. We thank Rika Maruyama for development of *tbc1* KO construct. We thank Michael Strayhorn for immunostaining the *Tsn* knockout and binding site mutants with the secretory marker. We thank other members of the Andrew lab for their support and ideas at various phases of this project.

References

Abrams, E.W., and Andrew, D.J. (2005). CrebA regulates secretory activity in the Drosophila salivary gland and epidermis. Development *132,* 2743-2758.

Andrew, D.J., Baig, A., Bhanot, P., Smolik, S.M., and Henderson, K.D. (1997). The Drosophila dCREB-A gene is required for dorsal/ventral patterning of the larval cuticle. Development *124,* 181-193.

Ao, J., Wei, C., Si, Y., Luo, C., Lv, W., Lin, Y., Cui, Y., and Gao, X. (2015). Tudor-SN Regulates Milk Synthesis and Proliferation of Bovine Mammary Epithelial Cells. International Journal of Molecular Sciences *16,*

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME Suite: tools for motif discovery and searching. Nucleic Acids Res. *37,* W202-W208.

Barbosa, S., Fasanella, G., Carreira, S., Llarena, M., Fox, R., Barreca, C., Andrew, D., and O'Hare, P. (2013). An orchestrated program regulating secretory pathway genes and cargos by the transmembrane transcription factor CREB-H. Traffic *14,* 382-398.

dit Frey, N.F., Muller, P., Jammes, F., Kizis, D., Leung, J., Perrot-Rechenmann, C., and Bianchi, M.W. (2010). The RNA Binding Protein Tudor-SN Is Essential for Stress Tolerance and Stabilizes Levels of Stress-Responsive mRNAs Encoding Secreted Proteins in *Arabidopsis*. Plant Cell *22,* 1575-1591.

Elbarbary, R.A., Miyoshi, K., Myers, J.R., Du, P., Ashton, J.M., Tian, B., and Maquat, L.E. (2017). Tudor-SN-mediated endonucleolytic decay of human cell microRNAs promotes G1/S phase transition. Science *356,* 859-862.

Fox, R.M., and Andrew, D.J. (2015). Changes in organelle position and epithelial architecture associated with loss of CrebA. Biology Open *4,* 317-330.

Fox, R.M., Hanlon, C.D., and Andrew, D.J. (2010). The CrebA/Creb3-like transcription factors are major and direct regulators of secretory capacity. The Journal of Cell Biology *191,* 479-492.

Fox, R.M., Vaishnavi, A., Maruyama, R., and Andrew, D.J. (2013). Organ-specific gene expression: the bHLH protein Sage provides tissue specificity to Drosophila FoxA. Development *140,* 2160-2171.

Frakes, A.E., and Dillin, A. (2017). The UPRER: Sensor and Coordinator of Organismal Homeostasis. Molecular Cell *66,* 761-771.

Gong, W.J., and Golic, K.G. (2004). Ends-out, or replacement, gene targeting in Drosphila. Proceedings of the National Academy of Sciences *100,* 2556-2561.

Gratz, S.J., Ukken, F.P., Rubinstein, C.D., Thiede, G., Donohue, L.K., Cummings, A.M., and O'Connor-Giles, K.M. (2014). Highly Specific and Efficient CRISPR/Cas9-Catalyzed Homology-Directed Repair in *Drosophila*. Genetics *196,* 961-971.

Gutierrez-Beltran, E., Denisenko, T.V., Zhivotovsky, B., and Bozhkov, P.V. (2016). Tudor staphylococcal nuclease: biochemistry and functions. Cell Death Differ. *23,* 1739-1748.

Henderson, K.D., and Andrew, D.J. (2000). Regulation and Function of Scr, exd, and hth in the Drosophila Salivary Gland. Dev. Biol. *217,* 362-374.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. *4,* 44-57.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. *37,* 1-13.

Iyer, S.C., Ramachandran Iyer, E.P., Meduri, R., Rubaharan, M., Kuntimaddi, A., Karamsetty, M., and Cox, D.N. (2013). Cut, via CrebA, transcriptionally regulates the COPII secretory pathway to direct dendrite development in Drosophila. J. Cell. Sci. *126,* 4732-4745.

Jennings, R.E., Berry, A.A., Strutt, J.P., Gerrard, D.T., and Hanley, N.A. (2015). Human pancreas development. Development *142,* 3126-3137.

Keller, R.B., Tran, T.T., Pyott, S.M., Pepin, M.G., Savarirayan, R., McGillivray, G., Nickerson, D.A., Bamshad, M.J., and Byers, P.H. (2017). Monoallelic and biallelic CREB3L1 variant causes mild and severe osteogenesis imperfecta, respectively. Genet. Med.

Kittler, R., Zhou, J., Hua, S., Ma, L., Liu, Y., Pendleton, E., Cheng, C., Gerstein, M., and White, K.P. (2013). A comprehensive nuclear receptor network for breast cancer cells. Cell. Rep. *3,* 538-551.

Kudron, M.M., Victorsen, A., Gevirtzman, L., Hillier, L.W., Fisher, W.W., Vafeados, D., Kirkey, M., Hammonds, A.S., Gersch, J., Ammouri, H.*, et al.* (2018). The ModERN Resource: Genome-Wide Binding Profiles for Hundreds of Drosophila and Caenorhabditis elegans Transcription Factors. Genetics *208,* 937-949.

Larigot, L., Juricek, L., Dairou, J., and Coumoul, X. (2018). AhR signaling pathways and regulatory functions. Biochim. Open *7,* 1-9.

Lehmann, R., and Tautz, D. (1994). Chapter 30 In Situ Hybridization to RNA. Methods in Cell Biology *44,* 575-598.

Liu, T., Ortiz, J.A., Taing, L., Meyer, C.A., Lee, B., Zhang, Y., Shin, H., Wong, S.S., Ma, J., Lei, Y.*, et al.* (2011). Cistrome: an integrative platform for transcriptional regulation studies. Genome Biol. *12,* R83.

Loganathan, R., Lee, J.S., Wells, M.B., Grevengoed, E., Slattery, M., and Andrew, D.J. (2016). Ribbon regulates morphogenesis of the Drosophila embryonic salivary gland through transcriptional activation and repression. Dev. Biol. *409,* 234-250.

Mendelson, C.R. (2000). Role of Transcription Factors in Fetal Lung Development and Surfactant Protein Gene Expression. Annu. Rev. Physiol. *62,* 875-915.

Milon, B., Sun, Y., Chang, W., Creasy, T., Mahurkar, A., Shetty, A., Nurminsky, D., and Nurminskaya, M. (2014). Map of open and closed chromatin domains in Drosophila genome. BMC Genomics *15,* 988-2164-15-988.

Murakami, T., Saito, A., Hino, S., Kondo, S., Kanemoto, S., Chihara, K., Sekiya, H., Tsumagari, K., Ochiai, K., Yoshinaga, K.*, et al.* (2009). Signalling mediated by the endoplasmic reticulum stress transducer OASIS is involved in bone formation. Nat. Cell Biol. *11,* 1205-1211.

Myat, M.M., and Andrew, D.J. (2002). Epithelial Tube Morphology Is Determined by the Polarized Growth and Delivery of Apical Membrane. Cell *111,* 879-891.

Negre, N., Brown, C.D., Ma, L., Bristow, C.A., Miller, S.W., Wagner, U., Kheradpour, P., Eaton, M.L., Loriaux, P., Sealfon, R.*, et al.* (2011). A cis-regulatory map of the Drosophila genome. Nature *471,* 527-531.

Nitta, K.R., Jolma, A., Yin, Y., Morgunova, E., Kivioja, T., Akhtar, J., Hens, K., Toivonen, J., Deplancke, B., Furlong, E.E.M., and Taipale, J. (2015). Conservation of transcription factor binding specificities across 600 million years of bilateria evolution. ELife *4,* e04837.

Noyes, M.B., Meng, X., Wakabayashi, A., Sinha, S., Brodsky, M.H., and Wolfe, S.A. (2008). A systematic characterization of factors that regulate Drosophila segmentation via a bacterial one-hybrid system. Nucleic Acids Res. *36,* 2547-2560.

Ober, E.A., and Lemaigre, F.P. (2018). Development of the liver: Insights into organ and tissue morphogenesis. Journal of Hepatology *68,* 1049-1062.

Reuter, R., and Scott, M.P. (1990). Expression and function of the homoeotic genes Antennapedia and Sex combs reduced in the embryonic midgut of Drosophila. Development *109,* 289-303.

Saito, A. (2014). Physiological functions of endoplasmic reticulum stress transducer OASIS in central nervous system. Anat. Sci. Int. *89,* 11-20.

Rose, R.E., Gallaher, N.M., Andrew, D.J., Goodman, R.H., and Smolik, S.M. (1997). The CRE-binding protein dCREB-A is required for Drosophila embryonic development. Genetics *146,* 595-606.

Saito, A., Hino, S., Murakami, T., Kanemoto, S., Kondo, S., Saitoh, M., Nishimura, R., Yoneda, T., Furuichi, T., Ikegawa, S.*, et al.* (2009). Regulation of endoplasmic reticulum stress response by a BBF2H7-mediated Sec23a pathway is essential for chondrogenesis. Nat. Cell Biol. *11,* 1197-1204.

Shin, J.J.H., Gillingham, A.K., Begum, F., Chadwick, J., and Munro, S. (2017). TBC1D23 is a bridging factor for endosomal vesicle capture by golgins at the trans-Golgi. Nature Cell Biology

Shinjo, S., Mizotani, Y., Tashiro, E., and Imoto, M. (2013). Comparative analysis of the expression patterns of UPR-target genes caused by UPR-inducing compounds. Biosci. Biotechnol. Biochem. *77,* 729-735.

Shintaku, J., Peterson, J., Talbert, E., Gu, J., Ladner, K., Williams, D., Mousavi, K., Wang, R., Sartorelli, V., and Guttridge, D. (2016). MyoD Regulates Skeletal Muscle Oxidative Metabolism Cooperatively with Alternative NF-κB. Cell Reports *17,* 514-526.

Slattery, M., Ma, L., Negre, N., White, K.P., and Mann, R.S. (2011). Genome-wide tissue-specific occupancy of the Hox protein Ultrabithorax and Hox cofactor Homothorax in Drosophila. PLoS One *6,* e14686.

Su, C., Zhang, C., Tecle, A., Fu, X., He, J., Song, J., Zhang, W., Sun, X., Ren, Y., Silvennoinen, O.*, et al.* (2015). Tudor staphylococcal nuclease (Tudor-SN), a novel regulator facilitating G1/S phase transition, acting as a co-activator of E2F-1 in cell cycle regulation. J. Biol. Chem. *290,* 7208-7220.

Troha, K., Im, J.H., Revah, J., Lazzaro, B.P., and Buchon, N. (2018). Comparative transcriptomics reveals CrebA as a novel regulator of infection tolerance in D. melanogaster. PLOS Pathogens *14,* e1006847.

van IJzendoorn, S.C.D. (2006). Recycling endosomes. J. Cell. Sci. *119,* 1679-1681.

Weigel, D., Bellen, H.J., Jurgens, G., and Jackle, H. (1989). Primordium specific requirement of the homeotic gene fork head in the developing gut of the Drosophila embryo. Roux's Archives of Developmental Biology *198,* 201-210.

Weiss, J.B., Suyama, K.L., Lee, H.H., and Scott, M.P. (2001). Jelly belly: a Drosophila LDL receptor repeat-containing signal required for mesoderm migration and differentiation. Cell *107,* 387-398

Yang, D., Jiang, T., Liu, J., Zhang, B., Lin, P., Chen, H., Zhou, D., Tang, K., Wang, A., and Jin, Y. (2018). CREB3 Regulatory Factor -mTOR-autophagy regulates goat endometrial function during early pregnancy. Biology of Reproduction ioy044-ioy044.

ZeRuth, G.T., Takeda, Y., and Jetten, A.M. (2013). The Kruppel-like protein Gli-similar 3 (Glis3) functions as a key regulator of insulin transcription. Mol. Endocrinol. *27,* 1692-1705.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9,* R137-2008-9-9-r137. Epub 2008 Sep 17.

CHAPTER 2

Drosophila Tbc1 is found in the trans portion of the *trans* Golgi Network and co-localizes to a subset of Rabs

Abstract

Tbc1d23, a predicted Rab-GAP, has been shown to bridge two Golgins to the WAVE complex in the mammalian trans Golgi network (TGN). Through this activity, Tbc1d23 is proposed to link endosomally-derived vesicles to their target membrane in the trans Golgi. In Chapter 1, we showed that the Drosophila homolog of Tbc1d23 - CG4552/Tbc1 - localizes to the trans Golgi, is transcriptionally regulated by the CrebA transcription factor, and that *tbc1* loss leads to secretion defects. Here, we show that loss of *tbc1* also leads to salivary gland apical membrane irregularities and that Tbc1 co-localizes with a subset of Rabs, including Rab5 and Rab 11. We also demonstrate that Tbc1 does not affect levels of Rab5 or Rab11.

Introduction

Many organs are organized as epithelial tubes. For example, our entire respiratory, circulatory and excretory systems are composed of epithelial tubes, as well as the digestive tract and secretory organs. Even the central nervous system begins as an epithelial tube (Colas and Schoenwolf, 2001). Epithelial tubes are characterized by an epithelial monolayer often enveloped by an outer extracellular matrix (ECM). In some epithelial organs, there is an outer layer of supporting cells, such as smooth muscle, and/or an inner (lumenal) ECM (Andrew and Ewald, 2010). Epithelial cells are connected to each other by adherens junctions, as well as tight junctions (vertebrates) or septate junctions (invertebrates), and have apical-basolateral polarity. Although the basic organization of epithelial tubes is shared, each tubular organ has unique architecture and distinct functions. Consider, for example, the differences between the aorta (a relatively simple closed tube) and the lungs (multi-lobed, highly branched tubes). The cellular and molecular events underlying formation of these different organs is difficult to study in mammals, even in the relatively tractable mouse. On the other hand, the Drosophila embryonic salivary gland (SG) provides an excellent model for studying the specification and formation of tubular organs. Drosophila is genetically tractable and, with over a hundred years of use as a model system, has many useful tools. Importantly, since many transcription factors and signaling pathways in Drosophila are found in mammals, what we learn in the fly SG is likely to be directly relevant to the formation of human tubular organs.

Although simple in form, the embryonic SG undergoes several morphological changes during its development. The glands begin as a pair of placodes of around 140 cells each on the ventral surface of the embryo (Maruyama and Andrew, 2012). These glands invaginate dorsally and form an incipient tube, which continues to elongate as more cells internalize. Once the salivary gland cells contact the overlying gut mesoderm, they turn and migrate posteriorly. The transition from a plate of cells on the surface to a fully internalized and correctly positioned tube

is mediated by cell shape changes, cell rearrangement and active migration, with no cell division or cell death once the cells are specified to form salivary glands.

The salivary gland is initially specified by three homeobox transcription factors: Sex combs reduced (Scr), Extradenticle (Exd) and Homothorax (Hth) (Henderson and Andrew, 2000). These three transcription factors, in turn, activate transcription of additional transcription factors, including CrebA. One of the downstream targets of CrebA binding and activation is *CG4552*, also known as *tbc1*. Loss of CrebA reduces *tbc1* SG expression to the background levels observed in surrounding non-SG tissues. Ectopic expression of CrebA in cells that normally express the *engrailed* gene results in the ectopic expression of *tbc1* in the same cells.

Some work has been done with Drosophila *tbc1* and its orthologues in other species. For example, *tbc1* was identified in a large RNAi screen in Drosophila S2 cells, where they showed that reduction of *tbc1* results in reduced levels of phagocytosis (Stroschein-Stevenson et al., 2005). Also, knockdown of *tbc1*in the border cells (BC) of the developing ovary using RNAi resulted in slower BC migration (Laflamme et al., 2012). A study looking at innate immunity in *C. elegans* and in vertebrate macrophages suggested a role for Tbc1 orthologues in the innate immune response (Alper et al., 2008). Further studies in the murine system revealed a role for Tbc1d23 attenuating the innate immune response after initiation. Interestingly, Tbc1d23 is down-regulated in macrophages infected by *P. carinii* (De Arras et al., 2012). A more recent study showed that Tbc1d23 functions in endosome to Golgi trafficking, linking golgin-97 and golgin-245 in the *trans* Golgi to the WASH complex on endosome derived vesicles (Shin et al., 2017). There is only a single Tbc1 orthologue in each of the higher eukaryotic species that have been studied and the orthologues have very high levels of sequence conservation, suggesting a conserved role. To that end, we used the Drosophila embryonic SG to characterize the role of Tbc1 in a developmental context. We localize Tbc1 to distinct subcellular domains and carry out co-localization experiments with YFP-tagged Rab proteins known to also localize to these domains. We demonstrate that loss or overexpression of Tbc1 does not affect the total levels of

two of the Rab proteins that colocalize with Tbc1. Finally, we show that loss of *tbc1* results in SG apical membrane irregularities.

Results

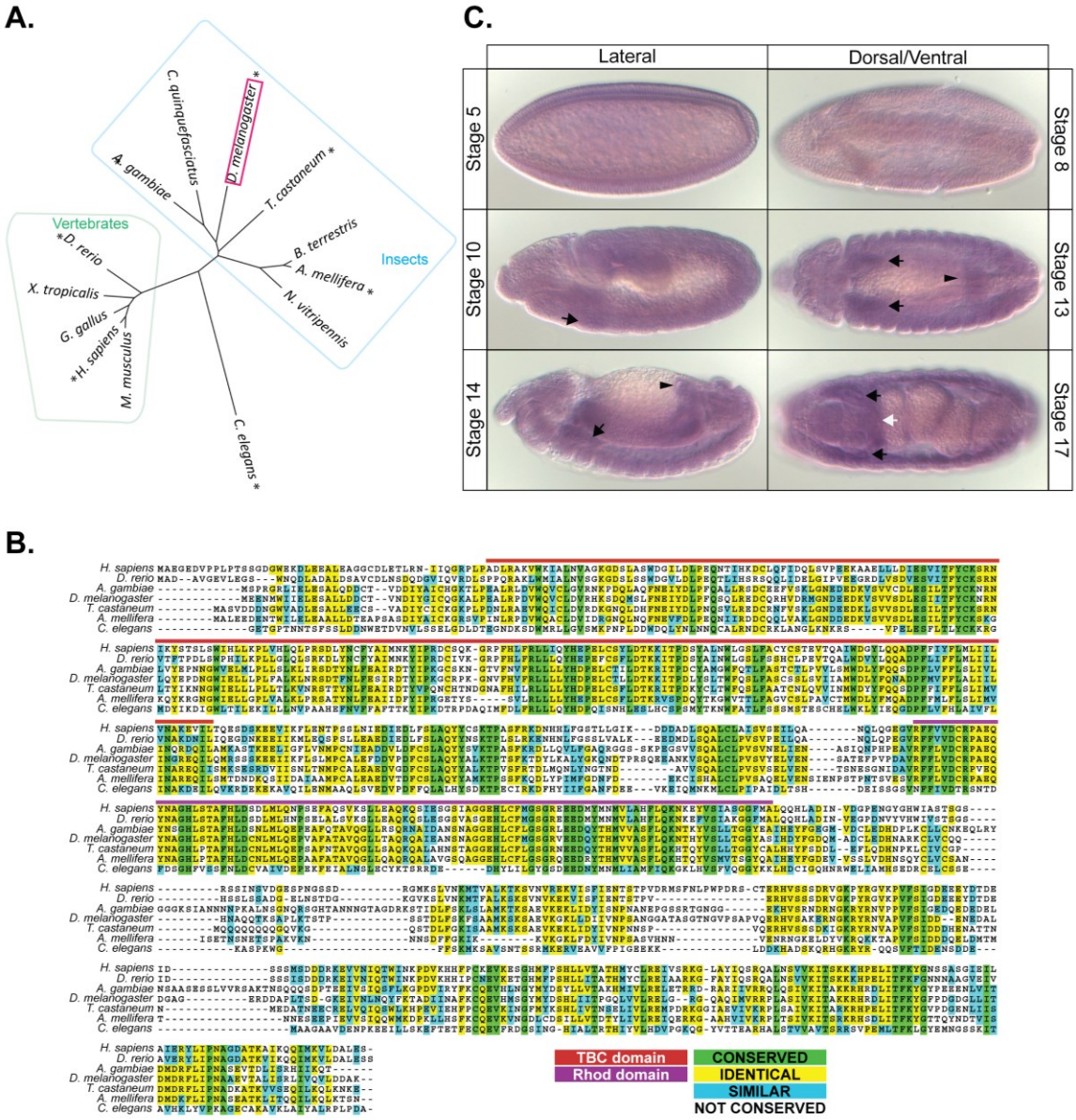
A BLASTp search with Drosophila Tbc1 identified a single orthologue in each species of higher eukaryotes (Fig. 19A). An alignment of the open reading frames (ORF) from a subset of these species revealed that the Tbc1 orthologs contain both a Tre1/Bub-2/Cdc-16p (TBC) domain and a Rhodanese-like domain (Fig. 19B). Tbc1 is the Drosophila representative of the TBC family branch I (Gabernet-Castello et al., 2013). The TBC domain of many family members has been shown to have Rab-GAP activity, but Tbc1 and its orthologues lack the key sequence motif known to be required for such activity. Thus, it is unclear if Tbc1 acts as a Rab-GAP. It has been recently shown that the mammalian orthologue acts as a bridging factor for two *trans* Golgi golgins and endosomally-derived vesicles.

The Rhodanese-like domain has several potential functions within a cell, including acting as a phosphatase, a ubiquitin hydrolase, and a sulfur-transferase (Cipollone et al., 2007). Additionally, the presence of the Rhodanese-like domain distinguishes Tbc1 and its orthologues from other TBC-containing proteins. There is high conservation throughout the protein among all orthologues, both within and outside of the TBC and Rhodanese-like domains. The existence of only a single orthologue per species, with very high levels of conservation across the animal kingdom, suggests that Tbc1 may have shared functions in all higher eukaryotes.

In Drosophila, *tbc1* transcripts are maternally contributed and transcripts are detectable at low levels in all cells of the developing embryo (Fig. 19C). Tbc1 is also transcriptionally upregulated in the SG, proventriculus, and hindgut. Upregulation in the SG begins early, when the SG is first specified, and continues throughout embryogenesis. Because Tbc1 is expressed in the Drosophila SG, we decided to use this tissue to investigate the normal physiological role of Tbc1.

A single P-element insertion line was available for *tbc1*, but the insertion was in an intron and had no effect on Tbc1 expression (data not shown). To generate a null mutation of *tbc1*, we used homologous recombination to replace the ORF with the *white*+ (w^+) eye color gene (Fig. 13). The *tbc1* null mutants were verified using PCR and *in situ* hybridization (Fig. 13). We

Figure 19: Drosophila Tbc1 is highly conserved and is expressed dynamically in embryos A) Unrooted tree of Tbc1 orthologues constructed by the Protein Tools (previously) available through the San Diego Supercomputer (http://workbench.sdsc.edu/). Asterisks indicate species included in the alignment. B) Alignment of Tbc1 orthologues. Red bar indicates the TBC domain, purple bar the Rhodanese-like domain. Green residues indicate conservation in all species. Yellow residues are the same many species and the cyan residues indicate similarity. C) mRNA accumulation pattern of *tbc1* in embryos. Whole-mount *in situs* with probes detecting *tbc1* mRNA in wild-type embryos. Black arrows, salivary gland; white arrows, proventriculus; arrowhead, hindgut. In stage 17, there is hindgut expression that is not visible in the focal plane shown.

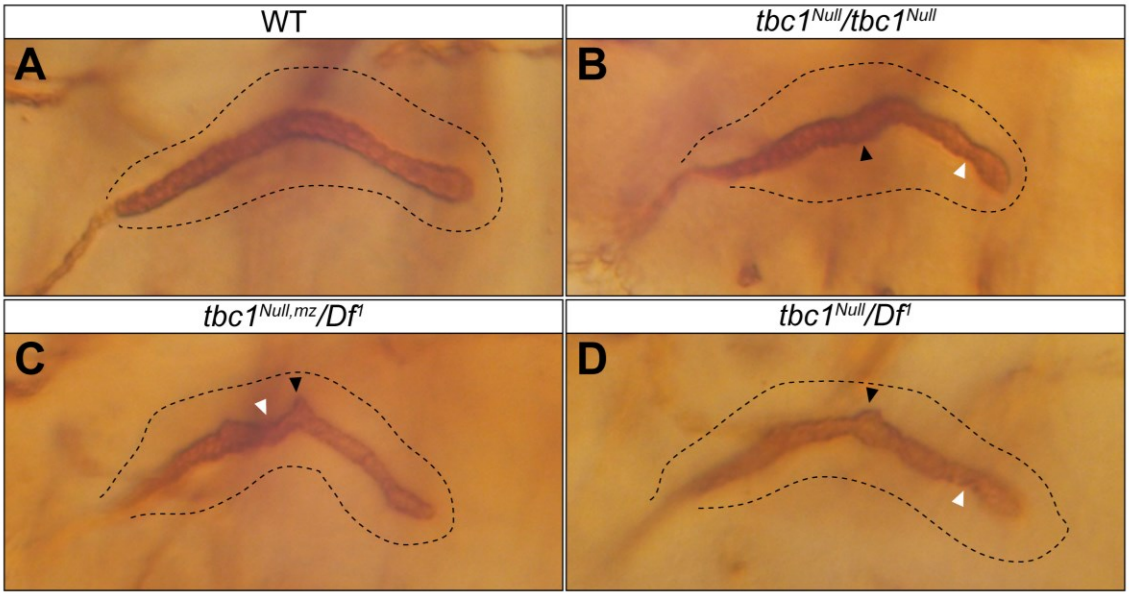

A.
C. quinquefasciatus
D. melanogaster *
A. gambiae
T. castaneum *
B. terrestris
A. mellifera *
N. vitripennis
Insects
Vertebrates
*D. rerio
X. tropicalis
G. gallus
*H. sapiens
M. musculus
C. elegans *
C.
Lateral
Dorsal/Ventral
Stage 5
Stage 8
Stage 10
Stage 13
Stage 14
Stage 17
B.
H. sapiens
D. rerio
A. gambiae
D. melanogaster
T. castaneum
A. mellifera
C. elegans
TBC domain
Rhod domain
CONSERVED
IDENTICAL
SIMILAR
NOT CONSERVED

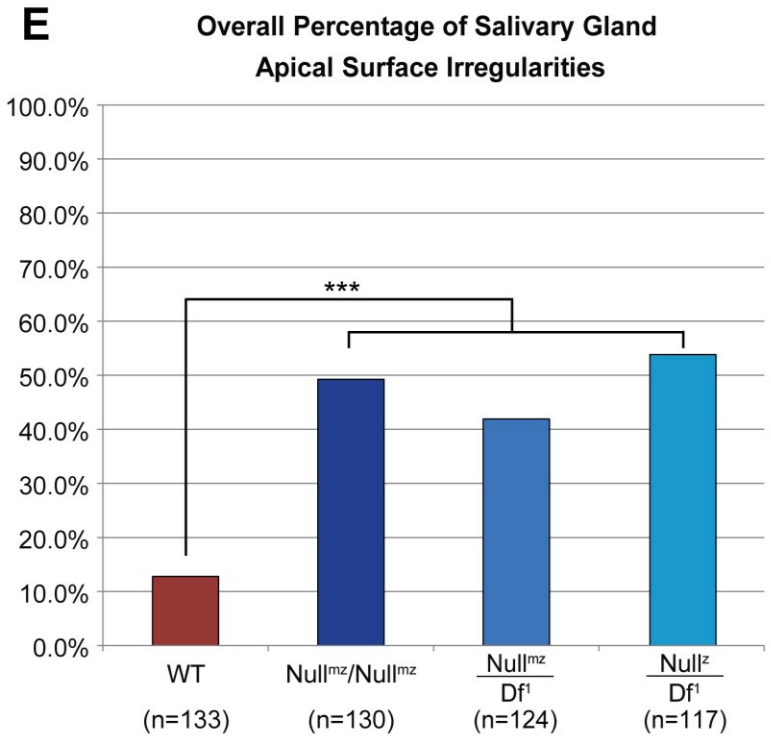
stained *tbc1* null embryos with an apical membrane marker and discovered that the SGs had between ~10% to ~50% apical membrane irregularities (Fig. 20). Trans-heterozygotes of the *tbc1* null and a deficiency that removes Tbc1 and several nearby genes had similar rates of apical membrane defects. Maternal loss of *tbc1* had no additional effect on the apical membrane defects, suggesting that it is the zygotic contribution of *tbc1* that is important for SG morphology. Egg chamber development and border cell migration were also examined, but no overt defects were noted (data not shown).

To determine where Tbc1 localizes within a cell, we overexpressed a C-terminal GFP-tagged version of Tbc1 in the SG using a Fkh-Gal4 driver. We stained for GFP and a variety of organellar markers. Tbc1 did not co-localize with the ER or with Csp, a marker for general secretory vesicles (Fig. 21A-B), but partially colocalized with recycling endosomes marked by Rab11 (Fig. 21C).

We subsequently developed a rat Tbc1 antiserum that detects Tbc1 with both immunohistochemistry and Western blotting (Fig. 22, 23). Using fluorescent confocal microscopy, we determined that endogenous Tbc1 stains in a punctate pattern in the cell, similar to the staining observed with GFP immunostaining of Tbc1-GFP expressed in the SG (Fig. 23). Staining with the Tbc1 antiserum in *tbc1* null embryos revealed low level non-specific staining near the SG apical membrane (Fig. 23A). Wild-type SGs have higher levels of apical membrane staining, suggesting that there is a pool of Tbc1 on or near the apical surface. As discussed in Chapter 1, endogenous staining of Tbc1 with Golgin-245 (*trans* Golgi marker) and GM130 (*cis* Golgi marker) showed Tbc1 did not colocalize with GM130 but partially colocalized with Golgin-245 (Fig. 14). In particular, there are multiple examples where we observe close staining of GM130 and Golgin-245, with close staining of Golgin-245 and Tbc1 (with a subset of Tbc1 overlapping with Golgin-245). This staining pattern is consistent with Tbc1 localizing to the trans most part of the TGN, where sorting of recycling endosomes occurs. We conclude that Tbc1 localizes to recycling endosomes and the TGN, suggesting that Tbc1 could be involved in the

Figure 20: Loss of *tbc1* results in irregular apical membrane surfaces in the salivary gland. A-D) Crb staining outlines the apical/lumenal surface. The salivary gland is outlined in each panel. Arrowheads indicate apical membrane irregularities. A) Wild-type, B) *tbc1* homozygous knockout, maternal-zygotic loss, C) Tbc1 knockout over a deficiency that removes *tbc1*, maternal-zygotic loss, D) zygotic loss of Tbc1 in knockout over the deficiency. E) Quantification of A-D; n indicates the number of salivary glands counted. ***: $p<0.000001$.


WT
A
tbc1Null/tbc1Null
B
tbc1Null,mz/Df1
C
tbc1Null/Df1
D


E
Overall Percentage of Salivary Gland Apical Surface Irregularities
100.0%
90.0%
80.0%
70.0%
60.0%
50.0%
40.0%
30.0%
20.0%
10.0%
0.0%
***
WT
(n=133)
Nullmz/Nullmz
(n=130)
Nullmz / Df1
(n=124)
Nullz / Df1
(n=117)

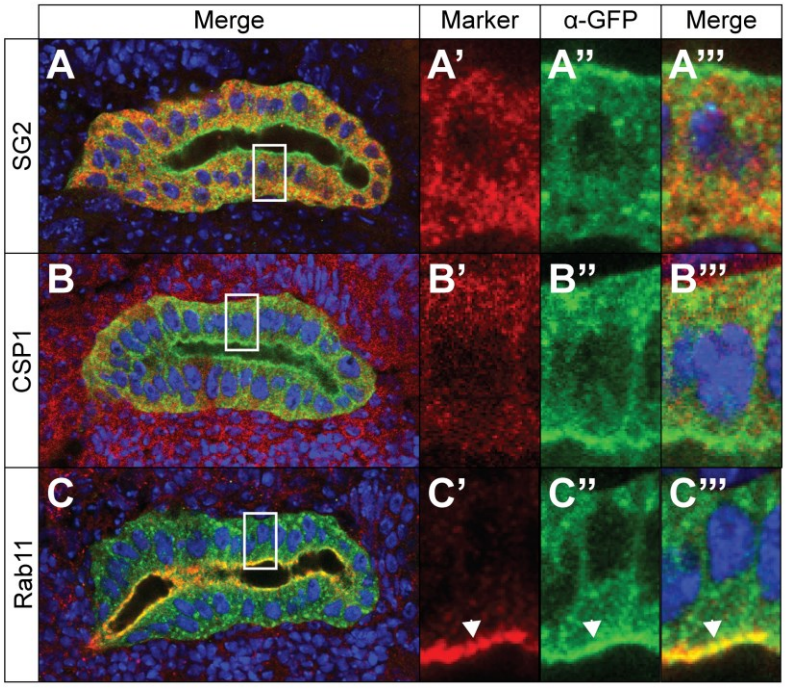
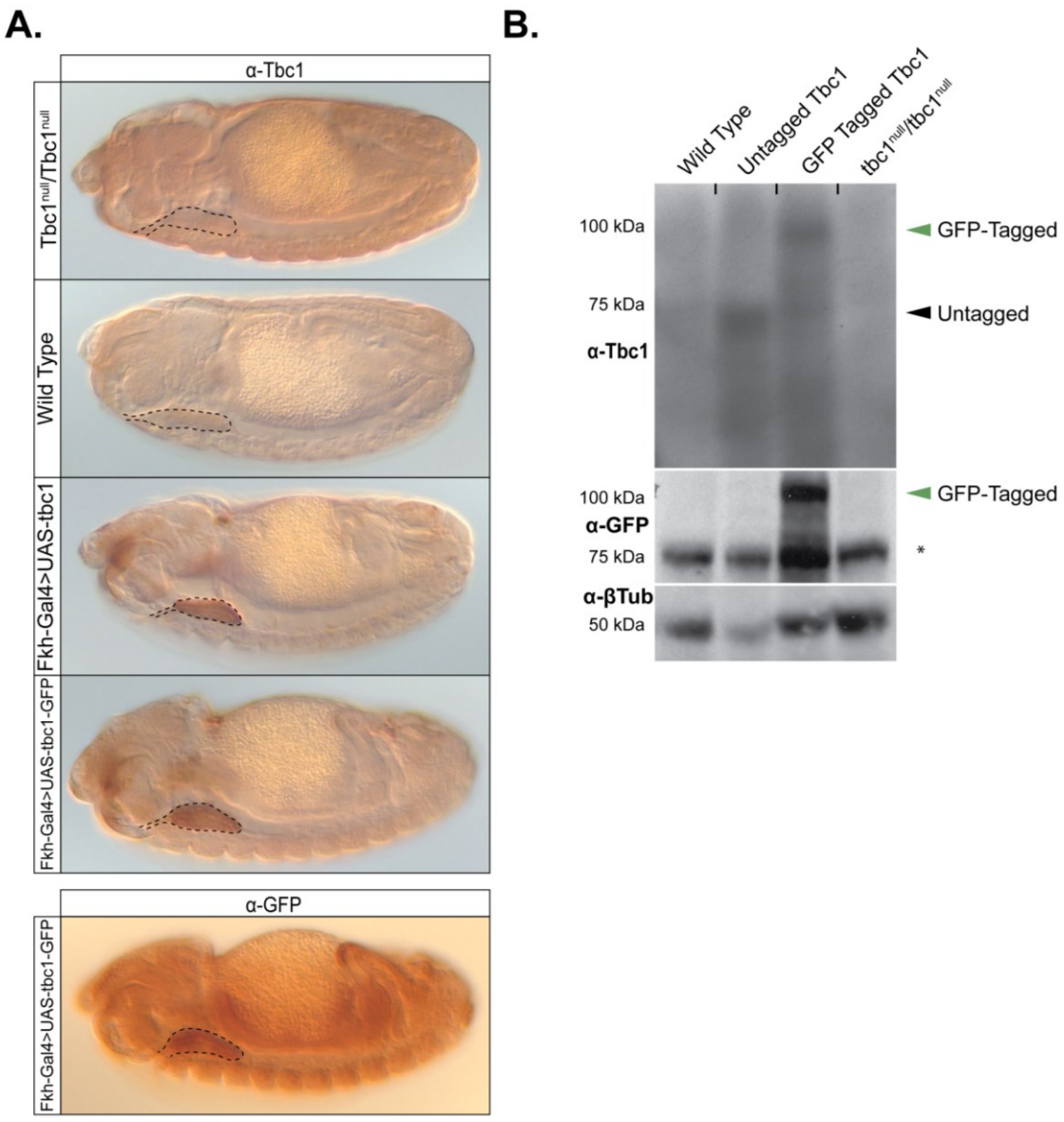


Figure 21: Tbc1 does not co-localize to dCSP or SG2 and partially colocalizes with Rab11. UAS-Tbc1-GFP driven by fkh-Gal4 and immuno-stained for GFP (Green). A) Costaining with the ER marker SG2 (Red). B) Costaining with the secretory vesicle marker dCSP1 (Red). C) Costaining with recycling endosomal marker Rab11 (Red). Blue: DAPI. White arrowhead: colocalization.

Figure 22: HRP and Western blotting using Tbc1 antiserum. A) Immunostaining with HRP. Staining is with the final bleed of Rat αTbc1 used at a dilution of 1:1000. Secondary was Rabbit α Rat-Biotin 1:500. B) HRP Immunohisochemistry and Western blotting with Tbc1 antiserum. A) HRP staining of embryos in $Tbc1^{Null}$, wild type and SG overexpression of untagged and GFP tagged Tbc1 using Tbc1 antiserum. SG is outlined with dashed line. B) HRP staining of GFP tagged Tbc1 overexpressed in the SG using GFP antibody. SG is outlined with dashed line. C) Western blotting for Tbc1 (top), GFP (middle) and βtub (bottom). Full length, untagged Tbc1: 77.5kDa; Full length, GFP tagged: 106.7kDa. Wild type: OR; Untagged: tub-Gal4>UAS-tbc1; Tagged: tub-Gal4>UAS-tbc1-GFP. Black arrowhead: Untagged Tbc1 size; Green arrowhead: expected size of GFP tagged Tbc1; *: non-specific bands.


A.
α-Tbc1
Tbc1null/Tbc1null
Wild Type
Fkh-Gal4>UAS-tbc1
Fkh-Gal4>UAS-tbc1-GFP
α-GFP
Fkh-Gal4>UAS-tbc1-GFP
B.
Wild Type
Untagged Tbc1
GFP Tagged Tbc1
tbc1null/tbc1null
100 kDa
75 kDa
α-Tbc1
GFP-Tagged
Untagged
100 kDa
α-GFP
75 kDa
GFP-Tagged
*
α-βTub
50 kDa

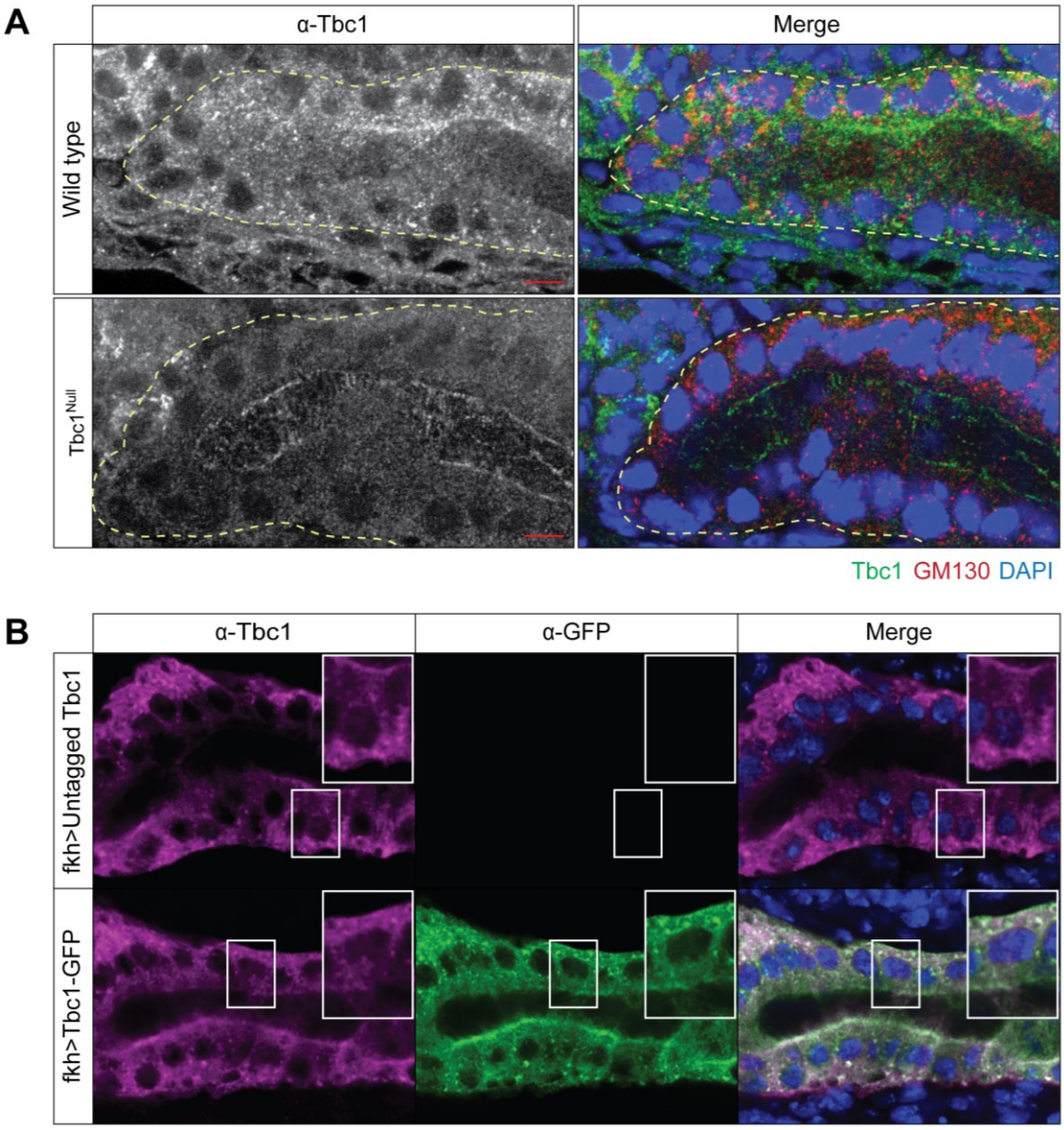


Figure 23: Fluorescent confocal images of using Tbc1 antibody. A) Tbc1 staining in wild type and $Tbc1^{Null}$. Tbc1 alone is shown in left hand panels and merge with DAPI (blue) and GM130 (red) on right. B) Tbc1 (purple) and GFP (green) staining of untagged (top) and GFP-tagged overexpressed Tbc1 in the SG. GFP and Tbc1 staining fully overlap in salivary glands overexpressing Tbc1-GFP. Blue: DAPI.

sorting of cargo destined for the apical membrane.

Because Tbc1 is a predicted Rab-GAP, we wanted to determine which, if any, Rabs might colocalize with Tbc1. Due to Tbc1's cytosolic localization, we focused on 19 of the more than 30 Rabs known to exist in flies. We expressed both untagged UAS-Tbc1 and UAS-YFP tagged forms of each of the Rabs in the SG using a fkh-Gal4 driver line (Zhang et al., 2007). For Rab7, we used UAS-Rab7-GFP. Of the 19 Rabs we tested, only a subset – Rab-X1, -5, -8, -9, and -10 – partially colocalized with Tbc1 (Fig. 24, 25).

Shin et al, 2017 showed that Tbc1d23 has a role besides the predicted Rab-GAP function, chiefly that it acts as a bridging factor in the TGN. Another possible role, suggested by the presence of the Rhodanese-like domain, is that Tbc1 could be involved in the degradation of interactors. The Rhodanese-like domain has many potential roles within the cell, one of which is hydrolyzing ubiquitin. Classically, ubiquitination can lead to degradation of targeted proteins. To ask if Tbc1 affected protein levels of Rab 5 or Rab11, we did Westerns and compared the levels of Rab5 and Rab11 in WT and *tbc1* null embryos (Fig. 26). We did not observe any overt differences in the levels of either Rab suggesting that Tbc1 is not involved in the synthesis or degradation of either Rab.

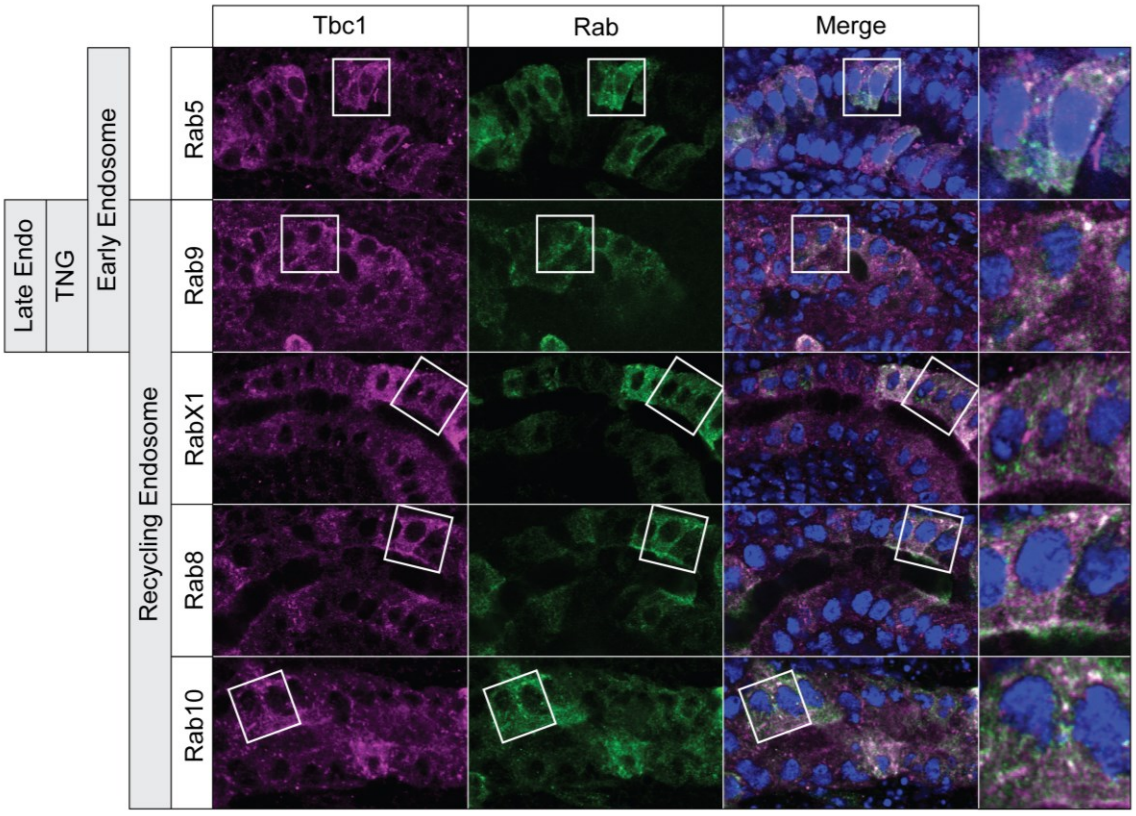


Figure 24: Partial Colocalization of Tbc1 with a subset of Rabs. Tbc1: purple; YFP-Rab: Green; DAPI: Blue. UAS-Tbc1 and UAS-YFP-Rab were overexpressed using a fkh-Gal4 driver on chromosome II that has mosaic expression in the SG.

Figure 25: Tbc11 does not colocalize with a subset of Rabs. Tbc1: purple; YFP-Rab or GFP-Rab: Green; DAPI: Blue. UAS-Tbc1 and UAS-YFP-Rab (or, in the case of Rab7 UAS-GFP-Rab7) were overexpressed using a fkh driver on chromosome II that has mosaic SG expression. *: non-specific binding of Tbc1.

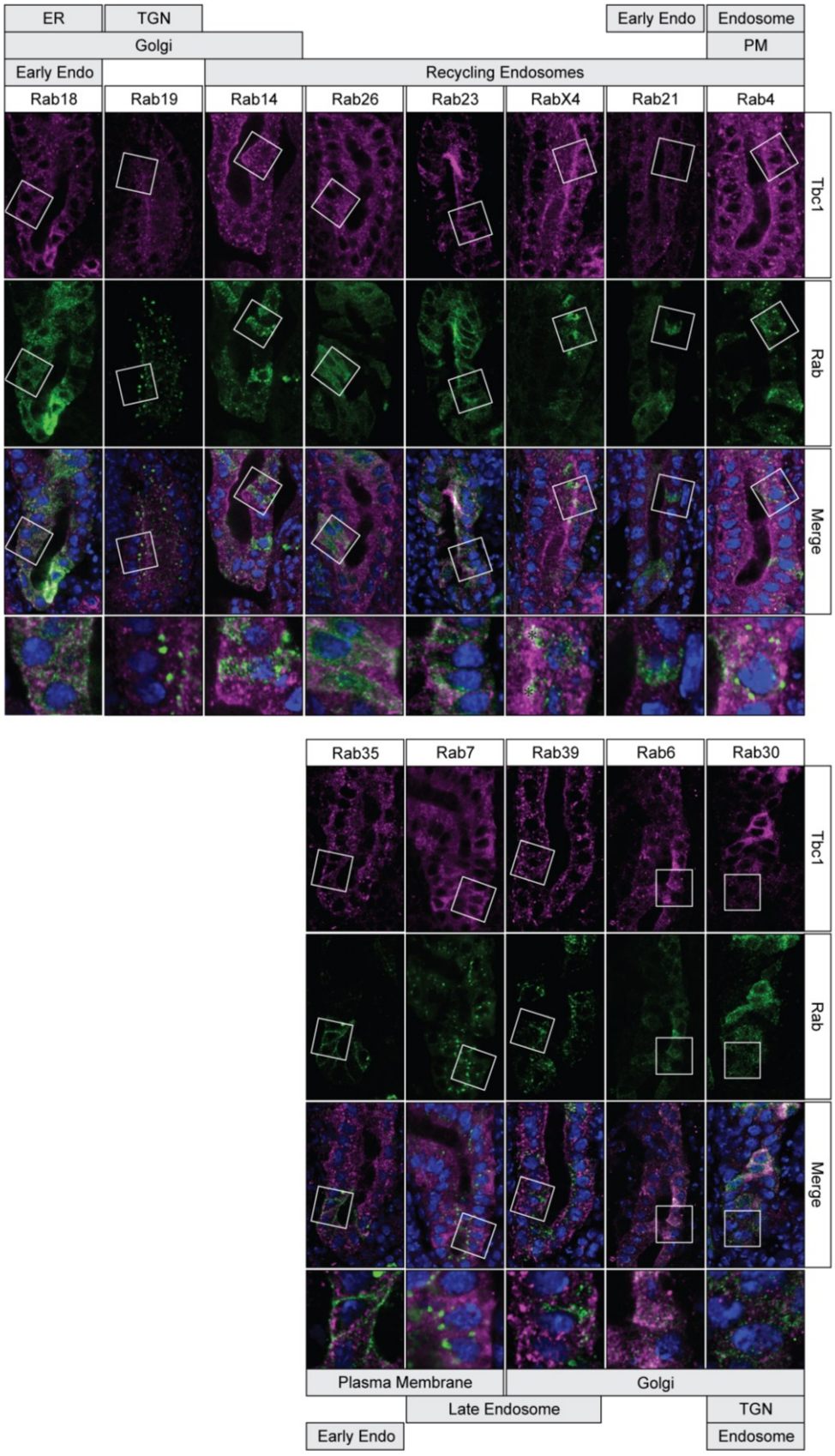
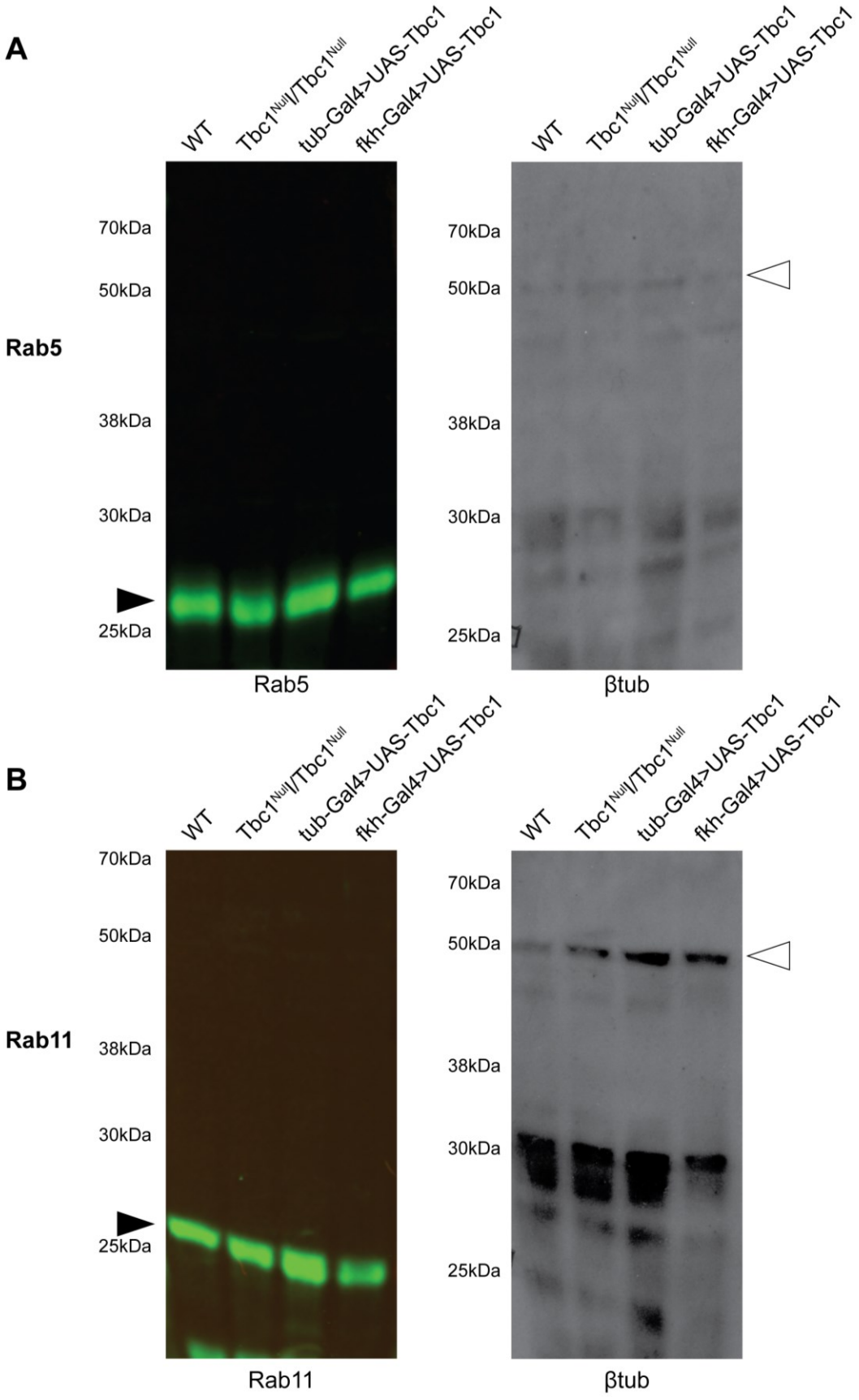

ER
TGN
Golgi
Early Endo
Early Endo
Endosome
PM
Recycling Endosomes
Rab18
Rab19
Rab14
Rab26
Rab23
RabX4
Rab21
Rab4
Tbc1
Rab
Merge
Rab35
Rab7
Rab39
Rab6
Rab30
Tbc1
Rab
Merge
Plasma Membrane
Golgi
Late Endosome
TGN
Early Endo
Endosome

Figure 26: Tbc1 loss or overexpression does not lead to changes in Rab5 or Rab11 levels. A) Rab5 with Western using LI-COR and βtub with chemiluminescence. B) Rab5 with Western using LI-COR and βtub with chemiluminescence Black arrowhead: Rab; White arrowhead: βtub.


A
WT
Tbc1Null/Tbc1Null
tub-Gal4>UAS-Tbc1
fkh-Gal4>UAS-Tbc1
70kDa
50kDa
Rab5
38kDa
30kDa
25kDa
Rab5
βtub
B
Rab11
Rab11
βtub

Discussion

Tbc1 has a clear orthologue in all higher eukaryotes (Fig. 18A) and the proteins are very highly conserved (Fig. 18B). The key characteristic of this family is the combination of the TBC domain (the Rab-GAP domain) and the Rhodanese-like domain. Due to the Rhodanese-like domain, there was a possibility that Tbc1 might prevent the degradation of interacting Rabs, as the domain can function as an ubiquitin hydrolase. We determined that neither Rab5 nor Rab11 protein levels changed in whole embryo lysates of wild type, *tbc1* loss, or *tbc1* overexpression (ubiquitously or SG-specific; Fig. 26), suggesting that Tbc1 does not play a role in the steady state levels of Rabs.

In Drosophila, Tbc1 is expressed ubiquitously but is also upregulated in multiple secretory tissues, such as the SG (Fig. 18C). SG expression of *tbc1* requires CrebA and CrebA is sufficient to induce ectopic *tbc1* mRNA expression (Fig. 8). Using a marker for SG secretion and cuticle preps, we show that loss of *tbc1* affects both SG and cuticle secretion. Since similar defects are observed with all other secretory pathway component gene (SPCG) mutations we have tested, these findings support a role for Tbc1 as an SPCG.

Recently, the mammalian homolog of Tbc1, Tbc1d23, was found to be a bridging factor in the TGN (Shin et al., 2017), connecting Golgins that localize to the trans Golgi with the WASH complex. Mutations in Tbc1d23 were also recently linked to Pontocerebellar hypoplasia (PCH) (Ivanova et al., 2017; Marin-Valencia et al., 2017). Ivanova et al (2017) showed that loss of Tbc1d23 in fibroblasts led to lysosomal trafficking and dense core vesicles defects. Additionally, in HeLa cells, Arl1 and Arl8 control Tbc11d23 levels at the TGN (Marin-Valencia et al., 2017). Like Tbc1d23, Tbc1 localizes to the trans-Golgi, specifically to the trans most portion of the TGN (Fig. 14). Using antibodies and YFP tags, we showed that Tbc1 partially colocalizes with six different Rabs (Fig. 21,24). Five of the six are known to localize to recycling endosomes. The one that is not, Rab5, has been localized to the early endosomes. Rab9 localizes to both recycling and early endosomes, as well as the TGN and late endosomes (Fig.24).

Altogether, this localization analysis supports the idea that Tbc1 and its homologs are involved in the targeting of vesicles to/from recycling endosomes and the TGN.

Other studies have hinted at an SPCG role of Tbc1 and its homologs. Using RNAi against Tbc1, Stroschein-Stevenson et al. (2005) showed that knockdown of *tbc1* led to a decrease in overall phagocytosis. One reason for this might be that cargo targeted to the phagocytosing membrane is mistargeted and/or slowed, preventing phagocytosis from occurring at normal rates. In *C. elegans* and mouse studies, loss of *tbc1* homologs led to longer innate immune responses to bacterial infection (Alper et al., 2008; De Arras et al., 2012). In tbc1 homolog knockout mice, it was found that the IL-6 cytokine was secreted at reduced levels and that the issue was not initialization of the innate immune response but rather the maintenance of high cytokine levels. A key part of the innate immune response in all animals is the secretion of antimicrobial peptides and/or cytokines. Loss of Tbc1 acting as bridging factor could lead to slower exocytosis. The TGN is central in targeting of vesicles to different organelles and areas of the plasma membrane and the loss of Tbc1 as a bridging factor could lead to mistargeting of vesicles and/or slower vesicular exit from the TGN, thus affecting secretion.

Materials and Methods

Fly Strains

Oregon R was used the wild-type strain. w; fkh-*Gal4* was developed previously in the Andrew lab (Henderson and Andrew, 2000). All crosses and collections were performed at 25ºC, unless noted otherwise. From Bloomington Stock Center (Indianapolis, IN) for balancing and line isolation: w; Sco/CyO, ftz-lacZ and w;; TM3/TM6B and Sp/ CyO, ftz-lacZ; Dr/TM6B, Ubx-*lacZ*. All YFP-Rab (Zhang et al., 2007) lines are noted in Table 2 and were obtained from Bloomington Stock Center (Indianapolis, IN).

Cloning

UAS-tbc1 and UAS-tbc1-GFP was cloned using the Gateway Cloning system. The *tbc1* ORF was subcloned from cDNA AT03044 into the pEnterD plasmid using TOPO cloning (Invitrogen). The *tbc1* ORF was then swapped into the pTW or pTWG (Untagged or GFP, respectively; Gateway Cloning System; Carnegie Institution) plasmid using LR Recombination (Invitrogen). The pTWG construct adds an in frame GFP ORF to the C-terminus of the *tbc1* ORF. Rainbow Transgenics, LLC injected this construct into w^{1118} embryos to create the transgenic insertions, which were mapped to specific chromosomes using a combination of chromosome-

Rab	Stock Number	Stock Genotype	Identifier	Citation
Rab10	9789	y1 w*; P{UASp-YFP.Rab10}21	RRID:BDSC_9789	Zhang et al 2007
Rab14	9793	y^1 w^*; P{UASp-YFP.Rab14}CG9967^{5L}	RRID:BDSC_9793	Zhang et al 2007
Rab18	9796	y1 w*; P{UAST-YFP.Rab18}CG977501	RRID:BDSC_9796	Zhang et al 2007
Rab19	24150	y1 w*; P{UAST-YFP.Rab19}Hr3902	RRID:BDSC_24150	Zhang et al 2007
Rab21	23242	y1 w*; P{UAST-YFP.Rab21}pog04	RRID:BDSC_23242	Zhang et al 2007
Rab23	9802	y[1] w[*]; ; P{w[+mC]=UASp-YFP.Rab23}01	RRID:BDSC_9802	Zhang et al 2007
Rab26	23245	y[1] w[*]; P{w[+mC]=UAST-YFP.Rab26}05/CyO	RRID:BDSC_23245	Zhang et al 2007
Rab30	9812	y1 w*; P{UASp-YFP.Rab30}Cdk410	RRID:BDSC_9812	Zhang et al 2007
Rab35	9821	y^1 w*; P{UASp-YFP.Rab35}15	RRID:BDSC_9821	Zhang et al 2007
Rab4	23269	y1 w*; P{UASp-YFP.Rab4}09	RRID:BDSC_23269	Zhang et al 2007
Rab5	24616	y^1 w^*; P{UASp-YFP.Rab5}02	RRID:BDSC_24616	Zhang et al 2007
Rab6	23251	y^1 w^*; P{UAST-YFP.Rab6}CG10082^{01}/CyO	RRID:BDSC_23251	Zhang et al 2007
Rab7	42705	w*; P{UAS-Rab7.GFP}2	RRID:BDSC_42705	Bellen, H. (2012.12.5). Bellen insertions.
Rab8	9782	y1 w*; P{UASp-YFP.Rab8}45	RRID:BDSC_9782	Zhang et al 2007
Rab9	9784	y1 w*; P{UASp-YFP.Rab9}22	RRID:BDSC_9784	Zhang et al 2007
RabX1	9840	y1 w*; P{UASp-YFP.RabX1}12	RRID:BDSC_9840	Zhang et al 2007
RabX4	9851	y1 w*; P{UASp-YFP.RabX4}how19	RRID:BDSC_9851	Zhang et al 2007

Table 2: GFP/YFP Tagged Rab stocks utilized.

specific balancer lines. See fly strains.

For bacterial induction, the C-terminal region of *tbc1* ORF (amino acids 321-689) from AT03044 was subcloned into pET15b using In-Fusion HD Cloning Kit (Invitrogen; primers in Supplemental Table 1).

Antibody Staining

Embryo fixation and immunohistochemistry were performed as previously described (Reuter et al., 1990). Antibodies, their concentration, source and use are found in Supplemental Table 2.

In situ hybridization

In situ hybridization was performed as previously described (Lehmann and Tautz, 1994). Antisense RNA probes were directed to the entire coding region of Tbc1.

Antiserum Generation

pET15b-tbc1 was transformed into BL21-DE3 cells. Induction occurred once the cells reached 0.77AU with 0.1M IPTG and cells were grown at 37°C for another 4 hours. Inclusion body prep was performed to isolate the C-terminal region of Tbc1 as described previously (Hanlon and Andrew, 2016). The resulting protein prep was sent to Covance (Denver, PA) for injection into a rat that had been prescreened for the absence of immuno-reactivity in Drosophila embryos.

Western Blotting

Western blotting was done as described previously (Ismat et al., 2013). 10% SDS-PAGE gels were used to size separate protein from whole embryo lysates. The size-separated proteins were then transferred to a polyvinylidene difluoride (PVDF) membrane and blotted as described

previously. If the blot was for LI-COR, the protocol was performed as described, only with the Odyssey Blocking Buffer (PBS) substituted. Licor Odyssey CLx was used for imaging the Licor blots. Antibodies used, along with concentration and source are found in Supplemental Table 2.

Microscopy

Confocal microscopy was performed on a Zeiss AxioObserver with 780-Quasar confocal module & FCS (NIH Grant S10 OD016374) or, Zeiss AxioObserver with LSM700 confocal module (NIH Grant S10 OD016374) or Zeiss Axiovert 200 with 510-Meta confocal module. All DIC microscopy was performed with a Zeiss Axiophot 2 with Janoptik ProgResC14 Plus opitical imaging system

Acknowledgements

Dr. Sean Munro for kindly providing Golgin-245 antibody prior to its availability from DSHB.

To Dr. Akira Nakamura for sending the Rab5 antibody.

References

Alper, S., Laws, R., Lackford, B., Boyd, W.A., Dunlap, P., Freedman, J.H., and Schwartz, D.A. (2008). Identification of innate immunity genes and pathways using a comparative genomics approach. Proceedings of the National Academy of Sciences *105,* 7016-7021.

Andrew, D.J., and Ewald, A.J. (2010). Morphogenesis of epithelial tubes: Insights into tube formation, elongation, and elaboration. Dev. Biol. *341,* 34-55.

Cipollone, R., Ascenzi, P., and Visca, P. (2007). Common themes and variations in the rhodanese superfamily. IUBMB Life *59,* 51-59.

Colas, J., and Schoenwolf, G.C. (2001). Towards a cellular and molecular understanding of neurulation. Developmental Dynamics *221,* 117-145.

De Arras, L., Yang, I.V., Lackford, B., Riches, D.W.H., Prekeris, R., Freedman, J.H., Schwartz, D.A., and Alper, S. (2012). Spatiotemporal Inhibition of Innate Immunity Signaling by the Tbc1d23 RAB-GAP. The Journal of Immunology *188,* 2905-2913.

Gabernet-Castello, C., O'Reilly, A.J., Dacks, J.B., Field, M.C., and Glick, B.S. (2013). Evolution of Tre-2/Bub2/Cdc16 (TBC) Rab GTPase-activating proteins. Mol. Biol. Cell *24,* 1574-1583.

Hanlon, C.D., and Andrew, D.J. (2016). Drosophila FoxL1 non-autonomously coordinates organ placement during embryonic development. Developmental Biology *419,* 273-284.

Henderson, K.D., and Andrew, D.J. (2000). Regulation and Function of Scr, exd, and hth in the Drosophila Salivary Gland. Dev. Biol. *217,* 362-374.

Ismat, A., Cheshire, A.M., and Andrew, D.J. (2013). The secreted AdamTS-A metalloprotease is required for collective cell migration. Development *140,* 1981-1993

Ivanova, E.L., Mau-Them, F.T., Riazuddin, S., Kahrizi, K., Laugel, V., Schaefer, E., de Saint Martin, A., Runge, K., Iqbal, Z., Spitz, M.*, et al.* (2017). Homozygous Truncating Variants in TBC1D23 Cause Pontocerebellar Hypoplasia and Alter Cortical Development. The American Journal of Human Genetics *101,* 428-440.

Laflamme, C., Assaker, G., Ramel, D., Dorn, J.F., She, D., Maddox, P.S., and Emery, G. (2012). Evi5 promotes collective cell migration through its Rab-GAP activity. The Journal of Cell Biology *198,* 57-67.

Lehmann, R., and Tautz, D. (1994). Chapter 30 In Situ Hybridization to RNA. Methods in Cell Biology *44,* 575-598.

Marin-Valencia, I., Gerondopoulos, A., Zaki, M.S., Ben-Omran, T., Almureikhi, M., Demir, E., Guemez-Gamboa, A., Gregor, A., Issa, M.Y., Appelhof, B.*, et al.* (2017). Homozygous Mutations in TBC1D23 Lead to a Non-degenerative Form of Pontocerebellar Hypoplasia. The American Journal of Human Genetics *101,* 441-450.

Maruyama, R., and Andrew, D.J. (2012). Drosophila as a model for epithelial tube formation. Developmental Dynamics *241,* 119-135.

Shin, J.J.H., Gillingham, A.K., Begum, F., Chadwick, J., and Munro, S. (2017). TBC1D23 is a bridging factor for endosomal vesicle capture by golgins at the trans-Golgi. Nature Cell Biology

Stroschein-Stevenson, S., Foley, E., O'Farrell, P.,H., and Johnson, A.D. (2005). Identification of *Drosophila* Gene Products Required for Phagocytosis of *Candida albicans*. PLoS Biol *4,* e4.

Zhang, J., Schulze, K.L., Hiesinger, P.R., Suyama, K., Wang, S., Fish, M., Acar, M., Hoskins, R.A., Bellen, H.J., and Scott, M.P. (2007). Thirty-One Flavors of Drosophila Rab Proteins. Genetics *176,* 1307-1322.

CHAPTER 3

The generation and testing of GFP tagged *fork head*, *Sage,* and *sens* for ChIP-seq

Abstract

Fkh, Sage, and Sens are transcription factors (TFs) that play key roles in the developing Drosophila embryonic salivary gland. All three are required for salivary gland survival and full levels of expression of salivary gland specific gene products. Fkh plays an additional role in regulating the process of morphogenesis – creating a three dimensional tube from the two dimensional salivary gland primordia. To determine how these proteins interact with chromatin and with each other to control salivary gland specific gene expression and to identify direct downstream target genes, we are expressing GFP tagged forms of these TFs in the embryonic SG and performing ChIP-seq using a GFP antibody. So far, we have generated lines expressing GFP-tagged version of all three TFS. We have shown that the GFP tagged TFs are fully functional; their overexpression phenocopies overexpression of untagged forms of the same proteins, and the GFP tagged TFs can rescue loss-of-function phenotypes. We have isolated chromatin and sent it to our collaborators at the University of Minnesota. We are awaiting the results.

Introduction

The Drosophila salivary gland is initially specified by three homeobox transcription factors: Sex combs reduced (Scr), Extradenticle (Exd) and Homothorax (Hth) (Fig 1) (Henderson and Andrew, 2000). These three transcription factors in turn activate transcription of four additional transcription factors: CrebA, Huckebein (Hkb), Sage, and Fork head (Fkh) (Maruyama et al., 2011). The salivary gland expression of Scr, Exd and Hth is transient; expression is turned off during early stages of salivary gland invagination (Henderson and Andrew, 2000). Thus, the role of Scr-Exd-Hth is to initiate salivary gland specification.

The four transcription factors activated by Scr, Exd and Hth are responsible for the morphology, survival, and function of the salivary gland. CrebA was discussed previously. The Sp1/Egr-like transcription factor Hkb, unlike all the other salivary gland transcription factors, is expressed only transiently in the salivary gland (Myat and Andrew, 2000a; Myat and Andrew, 2002) and its expression is turned off in the salivary gland shortly after invagination begins. Hkb's role in the salivary gland is in tube elongation. In *hkb* mutants, apical vesicular trafficking is defective, leading to salivary glands that are much shorter than in wild type. Overexpression of Hkb leads to the opposite phenotype of normally-elongated salivary glands with expanded apical lumens.

Sage, unlike all the other known transcription factors involved in salivary gland specification, is only expressed in the salivary gland (Chandrasekaran and Beckendorf, 2003). This basic helix-loop-helix transcription factor works with Fkh to control expression of salivary gland-specific genes. Indeed, Sage is what confers tissue specificity to Fkh in the salivary gland (Fox et al., 2013). One known target of Fkh/Sage is another transcription factor, known as Senseless (Sens) (Chandrasekaran and Beckendorf, 2003; Fox et al., 2013). Sens increases the expression levels of Fkh-Sage target genes(Fox et al., 2013). The targets of Sage, Fkh and Sens include tissue-specific secreted gene products and modifying enzymes, suggesting that these transcription factors work together to regulate high-level production of gene products unique to

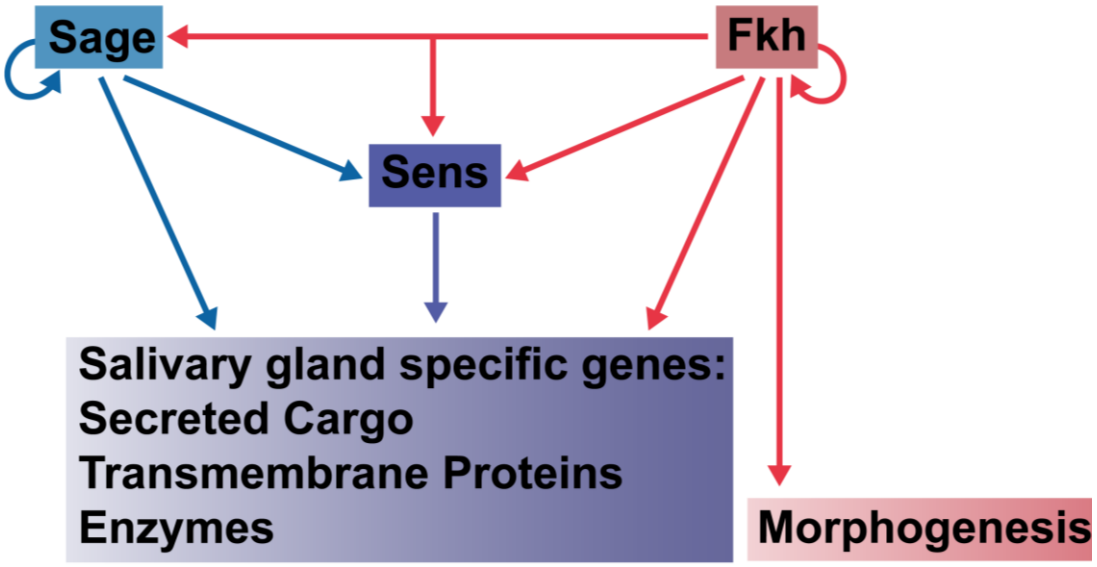


Figure 27: Fkh, Sage, and Sens pathway. modified from Fox et al, 2013.

the salivary gland. Fkh also appears to work with Sage and Sens to block apoptosis of SG cells, which cease dividing once they are specified (Chandrasekaran and Beckendorf, 2003; Fox et al., 2013; Myat and Andrew, 2000b).

Fkh has additional roles in salivary gland development that are not shared with Sage and Sens. Importantly, this includes driving the cell shape changes and cell rearrangements required to transform the two dimensional SG primordia into an elongated tube (Chung et al., 2017; Myat and Andrew, 2000b; Sanchez-Corrales et al., 2018). Interestingly, Fkh is also required to maintain expression of itself, CrebA and Sage after expression of Scr/Exd/Hth is turned off in the salivary gland. Although many downstream genes have been identified for Hkb, Sage, Fkh and Sens, the list of targets is almost certainly incomplete and whether known targets are directly regulated by Hkb, Sage, Fkh and Sens is unknown. Additionally, Fox et al 2013 hypothesized that Fkh, Sage and Sens act in a complex to activate expression of SG-specific genes and they showed that all three TFs bind to the same cytological locations on SG polytene chromosomes (Fig. 27).

To reveal the direct downstream targets of Fkh, Sage and Sens, all were tagged with GFP in preparation for salivary-gland specific ChIP-seq using the same protocol as was used with CrebA, which was described in Chapter 2. Further analysis of these lines showed that the GFP

tagged forms had similar defects when overexpressed as untagged versions of the proteins and they also could rescue the SG mutant phenotypes. Chromatin was prepared from lines expressing the GFP tagged versions of the three proteins expressed in the SG using two distinct Gal4 drivers that have in common expression in the embryonic SG. The chromatin was sent to our collaborator Matthew Slattery at The University of Minnesota, Duluth, and we are awaiting the ChIP-seq results.

Results

With the exception of Sage, all of the other known transcription factors that regulate salivary gland development are also expressed in other tissues. Consequently, we tagged *fkh*, *sage*, and *sens* with a C-terminal GFP and expressed the tagged forms in a tissue specific manner using *fkh-Gal4* and *sage-Gal4* (Fig. 28). Neither driver is SG specific but they have only expression in the SG in common. Using both drivers, we can determine the SG specific targets of the TFs as the intersection of the binding sites from both drivers. Additionally, since Sage is SG specific, we can perform ChIP-seq pulling down with a Sage antiserum developed in the Andrew lab using chromatin from wild-type embryos as a control.

To ensure the GFP tagged TFs are fully functional, we overexpressed both the untagged and GFP-tagged forms of each protein in an otherwise WT background and showed that the construct expressing the tagged versions result in the same phenotype as the untagged constructs (Fig. 29). For these experiments, a ubiquitous driver (*tub-Gal4*) was used and the embryos were stained for SG2 (an SG specific ER-resident protein). For *sage*, the overexpression embryos were fairly normal except that SG2 expression expanded into the SG duct and hindgut (Fig. 29A). This was observed previously with the untagged overexpression construct (Fox et al., 2013). Expression of Fkh or Sens led to gross morphological defects such as SGs that are not fully internalized and germband retraction defects (Fig. 29B, C)

To further test the functionality of the tagged versions of these transcription factors, we have rescued the loss-of-function phenotypes of *fkh*, sage, and *sens* (Fig. 30). Loss of any of these transcription factors leads to death of the SG cells. This is observed quite early with loss of Fkh (Fig. 30A), where the entire SG is lost at by embryonic stage 12 and is fully rescued by Fkh-GFP. When *sage* is knocked out, SG cells undergo massive apoptotic death at a slightly later embryonic stage (Fig. 30B). SG expression of Sage-GFP leads to full rescue of SG viability. The SG of a *sens* mutant embryo also undergoes late apoptosis resulting in a much smaller and lightly stained SG (Fig. 30C); the SG is the SG is completely normal when sens-GFP is expressed in the SGs of

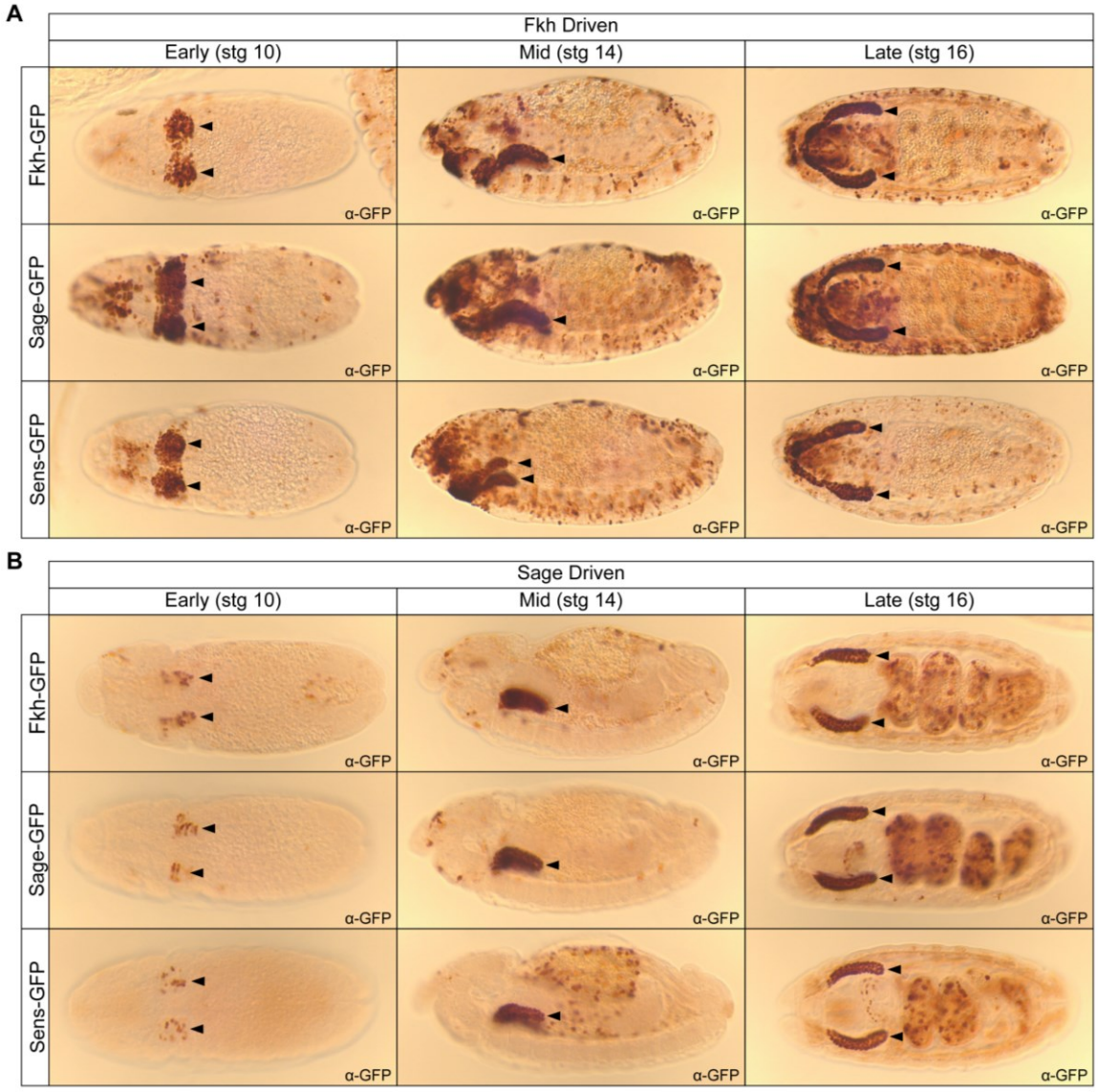


Figure 28: anti-GFP staining of all three constructs using Fkh-Gal4 (A) and Sage-Gal4 (B).

A) Fkh-Gal4: can see expression of GFP in the SG, hemocytes and peripheral nervous system. B) Sage-Gal4: expression of GFP in the SG and subset of gut cells. Black arrowheads: SG

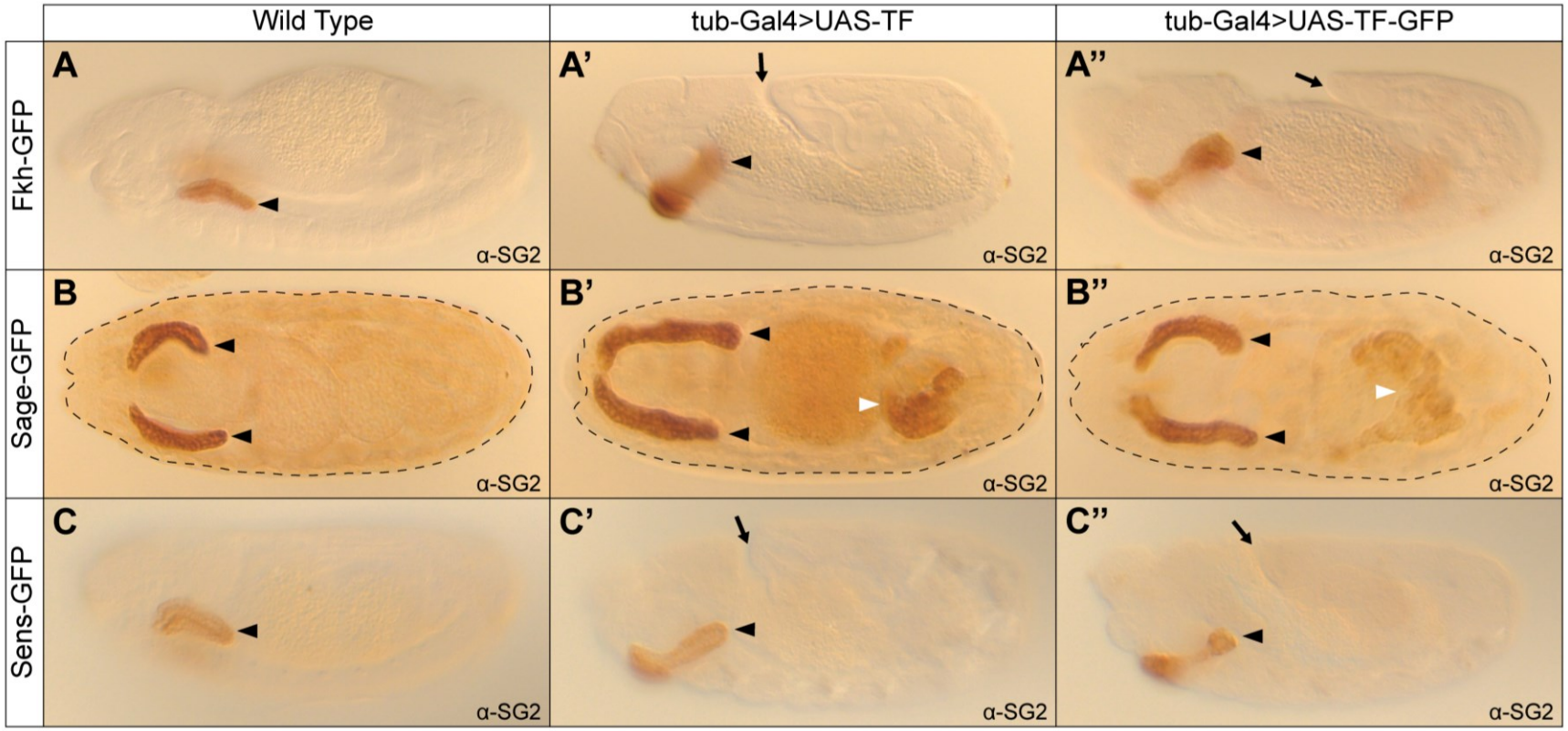
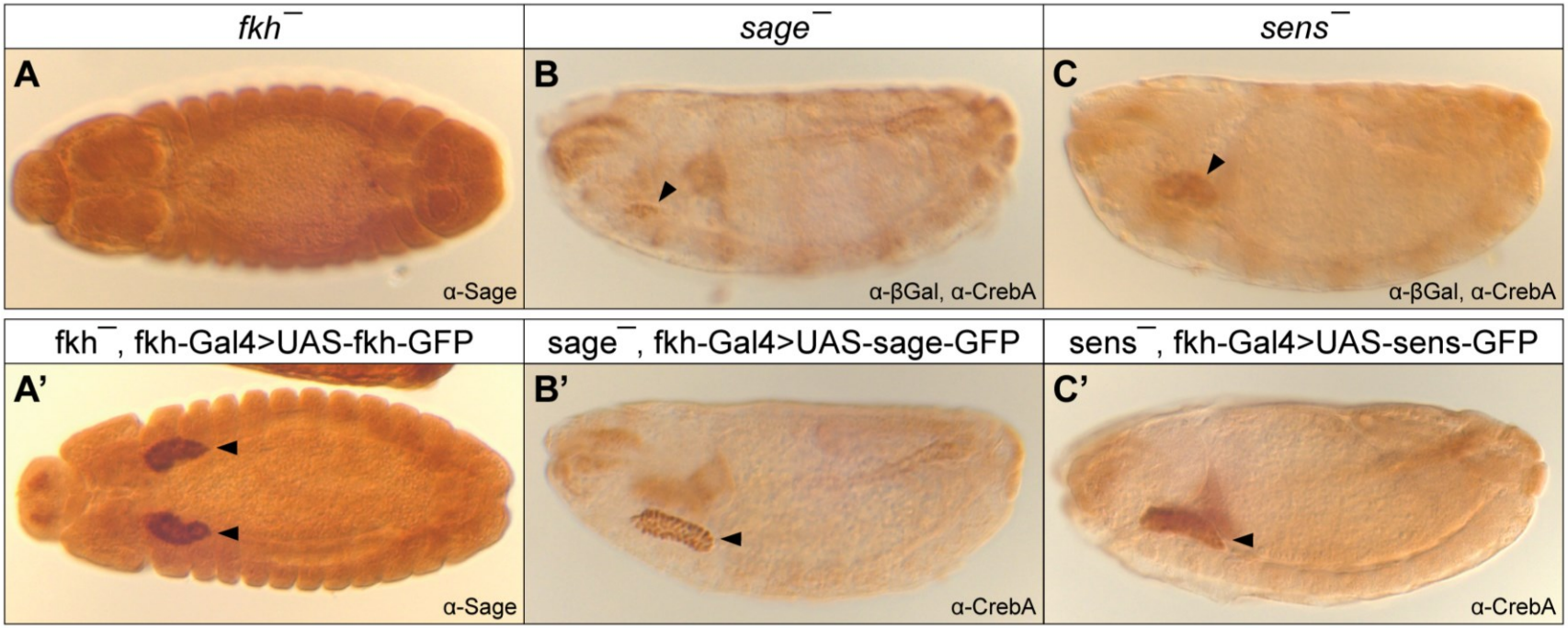


Figure 29: Tub-Gal4 overexpression of untagged and GFP-tagged Fkh, Sage, and Sens. A) tub-Gal4>UAS-fkh-GFP overexpression, there is a failure of germband retraction and SG internalization. B) tub-Gal4>UAS-sage-GFP overexpression, SG2 expression is expanded to duct and hindgut cells. C) tub-Gal4>UAS-sens-GFP overexpression, germband retraction and SG internalization has failed. Black arrowhead: SG; White arrowheads: the ectopic expression of SG2 in the hindgut. Black arrows: failure of germband retraction.



107

Figure 30: SG rescue of TF mutant phenotypes by SG-specific expression of GFP tagged TFs. A) Sage staining of *fkh* mutant and (‘) rescues. B and C) CrebA staining of B) *sage*, and C) *sens* mutants and (‘) rescues. *sage* and *sens* mutants are also stained with β-Gal and are siblings of the rescues. Black arrowheads indicate the SG cells.

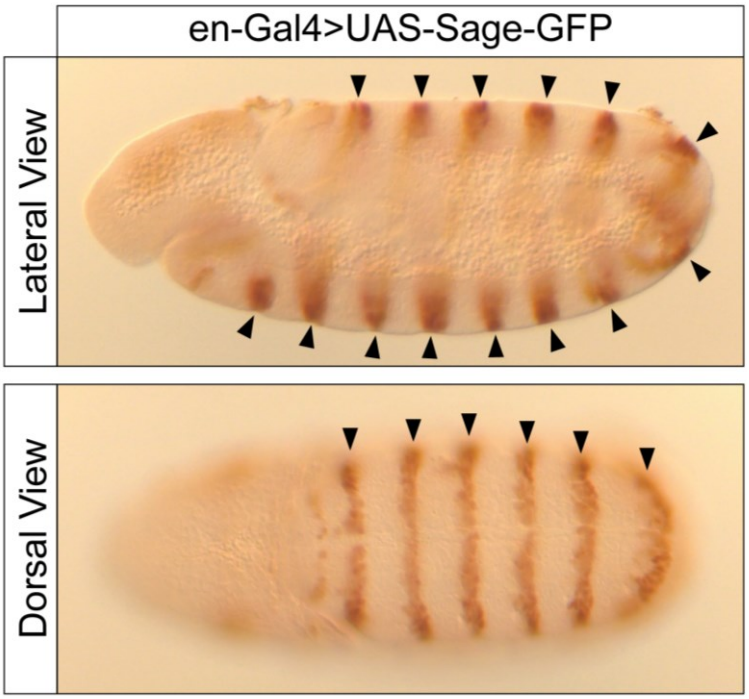


Figure 31: anti-Sage staining of en-Gal4 driven Sage-GFP. Arrows indicate expression in the en-Gal4 domains. SG has not been specified at this stage.

sens mutants. Thus, both the gain-of-function assays and the rescue of SG defects demonstrate that the GFP tagged form of *fkh*, *sage*, and *sens* are fully functional.

We also tested whether the Sage antisera recognizes Sage-GFP by overexpressing it using an *en-Gal4* line, which drives expression of transgenes containing a UAS promoter in stripes along the embryo (Figure 31). Conveniently, this antisera was previously used for ChIP-qPCR (Fox et al., 2013), suggesting that it will work well for the ChIP-seq experiments.

Three independent chromatin samples per genotypes have been collected from embryos expressing the GFP-tagged version of all three proteins under the control of two Gal4 lines that have only salivary gland expression in common. We have also obtained three independent chromatin samples from WT embryos, which will be immuno-precipitated for ChIP-Seq using the Sage antiserum, serving as an independent control for the Sage-GFP binding. All of the chromatin samples and the Sage antiserum was shipped to our collaborators at the University of Minnesota, Duluth. We are currently waiting for them to complete the ChIP-Seq analysis and send us the resulting data to continue our analysis.

Discussion

Future work will include ChIP-seq analysis using a well-tested GFP antibody to isolate chromatin from embryos expressing GFP tagged forms of three TFs: Fkh, Sage, and Sens, key TFs involved in the survival and function of the SG. Fkh has additional roles in SG morphogenesis. For our experiments, two Gal4 drivers that have only SG expression in common were used, thus enriching for SG-specific DNA binding events (Fig. 27). We have shown that the GFP tagged versions of each protein result in the same phenotypes when expressed ectopically as observed with the corresponding untagged versions (Fig. 28). We also showed that the GFP-tagged versions can rescue the SG specific phenotypes of the respective loss-of-function mutants (Fig. 29). Finally, with Sage we are able to show that the Sage antibody we intend to use as a control detects Sage-GFP when it is ectopically expressed in en stripes (Fig. 30). We have sent the chromatin to our collaborator at the University of Minnesota and are awaiting the resulting data from the ChIP-seq.

When comparing our GFP ChIP-seq data sets to existing published datasets (Negre et al 2011), we generally expect only a limited group of genes will be found in both. This is because we expect the other datasets to include genes under the control of the transcription factor in tissues other than the salivary gland. An exception to this rule is the dataset we generate using our Sage antisera. For the Sage datasets, we expect to find a one to one relationship since Sage is expressed in only the salivary gland.

Comparing our GFP datasets amongst themselves we expect to see binding overlap at some loci as it has been hypothesized that Fkh, Sage, and Sens work in a complex. Using known binding site motifs and analysis of the binding peaks we may determine the binding site requirements for the complex or individual TF to regulate downstream targets. We do expect that a subset of genes will not be bound by the complex but the individual TFs. Particularly because it is known that Fkh plays a separate role in the morphogenesis of the SG.

When we compare our data with existing microarray data (Fox et al, 2013; Maruyama et al, 2011; Liu and Lehmann, 2008), we expect to find many shared genes. The direction of the fold changes in the microarray will indicate if the transcription factor acts as an activator or repressor for each target. We will conclude that genes that are found to be targets in the microarray but not the ChIP-seq will be indirect targets of the transcription factor. Candidate target genes from the ChIP-seq not found in the microarray sets will be tested for regulation by performing similar *in situ* experiments as in Chapter 1. mRNA expression in wild-type embryos will be compared to mRNA expression in TF mutant embryos and embryos with overexpression of TF-GFP in en stripes. We may discover that although the transcription factor binds the DNA in the salivary gland, that regulation may be in other tissues because of additional requirements for cell-type specific regulators.

Materials and Methods

Fly Strains

Oregon R was used the wild-type strain. *w^{1118};UAS-fkh*, *w^{1118};UAS-sage*, *w^{1118};UAS-sens*, *w^{1118};fkh-Gal4*, and *w^{1118};Sage-Gal4/CyO, ftz-lacZ* were generated previously in the Andrew lab (Abrams et al., 2006; Fox et al., 2013; Henderson and Andrew, 2000). All crosses and collections were performed at 25ºC, unless noted otherwise. The *w;en-Gal4* is from Weiss et al., (2001). The following stocks were used for balancing transgenic and mutant lines: *w;Sco/CyO,ftz-lacZ* and *w;TM3/TM6B,Ubx-lacZ* and *Sp/CyO,ftz-lacZ; Dr/TM6B*,*Ubx-lacZ*.

Cloning

UAS-fkh-GFP, UAS-sage-GFP, and UAS-sens-GFP were cloned using the Gateway Cloning system. The *fkh* ORF and *sage* ORF were subcloned from cDNA RE03865 and RE59356, respectively, into the pEnterD plasmid using TOPO cloning (Invitrogen). UAS-sens (Nolo et al., 2000) flies were used as a template to amplify the ORF using primers: CS00 and CS01 (Table X). Site directed mutagenesis using the QuikChange kit (Aligent) was used to fix mutations in the ORF of *fkh*: T302C, G694A, and insertion of C into position 1274 (primer sequences are in Supplemental Table X). The corrected *fkh*, *sage*, and *sens* ORFs were then swapped into the pTWG (Gateway Cloning System; Carnegie Institution) plasmid using LR Recombination (Invitrogen); this construct adds an in frame GFP ORF to the C-terminus of the TF ORF. Rainbow Transgenics, LLC injected each construct into *w^{1118}* embryos to create the transgenic insertions, which were mapped to specific chromosomes using a combination of chromosome-specific balancer lines. See fly strains.

Antibody Staining

Embryo fixation and immunohistochemistry were performed as previously described (Reuter and Scott, 1990). All antisera used in this analysis, as well as the concentration and source of the antisera are listed in Supplemental Tables 2 and 3. All DIC microscopy was performed on a Zeiss Axiophot 2 light microscope equipped with a CCD camera.

Chromatin Preparation

Embryo collection and chromatin preparation were performed as described previously (Loganathan et al., 2016). Briefly, embryos were collected overnight on apple juice caps at 25ºC, rinsed and fixed in 1.8% formaldehyde for 15 min in a 1:3 solution with heptane. The liquid was removed and the fixed embryos were then stored at -80ºC until the chromatin preparation. Embryos were lysed during thawing using a dounce in a 1.8% formaldehyde solution. Chromatin was fragmented by sonication and separated from cellular debris through serial centrifugation. Once isolated, chromatin was stored at -80ºC and was shipped on dry ice.

Acknowledgements

Andrew Lab members for their help, especially SangJoon Kim for his excellent comments. Cierra Smith who was instrumental in generating the GFP tagged TFs.

References

Abrams, E.W., Mihoulides, W.K., and Andrew, D.J. (2006). Fork head and Sage maintain a uniform and patent salivary gland lumen through regulation of two downstream target genes, PH4aSG1 and PH4aSG2. Development *133,* 3517-3527.

Chandrasekaran, V., and Beckendorf, S.K. (2003). senseless is necessary for the survival of embryonic salivary glands in Drosophila. Development *130,* 4719-4728.

Chung, S., Kim, S., and Andrew, D.J. (2017). Uncoupling apical constriction from tissue invagination. ELife *6,* e22235.

Fox, R.M., Vaishnavi, A., Maruyama, R., and Andrew, D.J. (2013). Organ-specific gene expression: the bHLH protein Sage provides tissue specificity to Drosophila FoxA. Development *140,* 2160-2171.

Henderson, K.D., and Andrew, D.J. (2000). Regulation and Function of Scr, exd, and hth in the Drosophila Salivary Gland. Dev. Biol. *217,* 362-374.

Liu, Y., and Lehmann, M. (2008). A genomic response to the yeast transcription factor GAL4 in Drosophila. Fly (Austin) *2,* 92-98.

Loganathan, R., Lee, J.S., Wells, M.B., Grevengoed, E., Slattery, M., and Andrew, D.J. (2016). Ribbon regulates morphogenesis of the Drosophila embryonic salivary gland through transcriptional activation and repression. Dev. Biol. *409,* 234-250.

Maruyama, R., Grevengoed, E., Stempniewicz, P., and Andrew, D.J. (2011). Genome-Wide Analysis Reveals a Major Role in Cell Fate Maintenance and an Unexpected Role in Endoreduplication for the *Drosophila* FoxA Gene *Fork Head*. PLoS ONE *6,* e20901.

Myat, M.M., and Andrew, D.J. (2000b). Organ shape in the Drosophila salivary gland is controlled by regulated, sequential internalization of the primordia. Development *127,* 679-691.

Myat, M.M., and Andrew, D.J. (2000a). Fork head prevents apoptosis and promotes cell shape change during formation of the Drosophila salivary glands. Development *127,* 4217-4226.

Myat, M.M., and Andrew, D.J. (2002). Epithelial Tube Morphology Is Determined by the Polarized Growth and Delivery of Apical Membrane. Cell *111,* 879-891.

Negre, N., Brown, C.D., Ma, L., Bristow, C.A., Miller, S.W., Wagner, U., Kheradpour, P., Eaton, M.L., Loriaux, P., Sealfon, R.*, et al.* (2011). A cis-regulatory map of the Drosophila genome. Nature *471,* 527-531

Nolo, R., Abbott, L.A., and Bellen, H.J. (2000). Senseless, a Zn Finger Transcription Factor, Is Necessary and Sufficient for Sensory Organ Development in Drosophila. Cell *102,* 349-362.

Reuter, R., and Scott, M.P. (1990). Expression and function of the homoeotic genes Antennapedia and Sex combs reduced in the embryonic midgut of Drosophila. Development *109,* 289-303.

Sanchez-Corrales, Y., Blanchard, G.B., and Röper, K. (2018). Radially-patterned cell behaviours during tube budding from an epithelium. ELife *7,* e35717.

APPENDIX A

Supplemental Table 1

CrebA primers		
Primer Name	Primer Sequence	Use of primer
DJ052	GTACCTCACCTTCCACGTCCCGCCcACGCACGCCACGCCCATC	Site directed repair mutagenesis of CrebA; Forward
DJ053	GATGGGCGTGGCGTGCGTgGGCGGGACGTGGAAGGTGAGGTAC	Site directed repair mutagenesis of CrebA; Reverse
CS02	cacc ATGGAATTCTACGATGGCGACC	Cloning of UAS-CrebA-GFP; Forward
CS03	GGACTTTTTCACATTGTGCTTGC	Cloning of UAS-CrebA-GFP; Reverse
Xbp1 primers		
Primer Name	Primer Sequence	Use of primer
2R:21143405	TAAGATATTGCCTGAAATGTTAC	HR amplification primers for *Xbp1* ; forward
2R:21143758	CGTCTCATGAACTTCGACTGC	HR amplification primers for *Xbp1* ; reverse
Xbp1 5' Mut primer	CCAACTCTTCTAGAAAAACgttgtAGCTTTTCACAGCCATTCG	Mutagenesis primer for *Xbp1*
Xbp1 5' Mut primer	GCTTTTCACAGCCATTCGGAAAAGAACgttgtAGCAGAAAGGAAAACTGGTC	Mutagenesis primer for *Xbp1*
2R:21143639	CTTCGTCGGTCGATATTGCAAAGG	gRNA oligo cloning sequence *Xbp1* forward
2R:21143653	AAACCCTTTGCAATATCGACCGAC	gRNA oligo cloning sequence *Xbp1* reverse
TSN primers		
Primer Name	Primer Sequence	Use of primer
3L:264146	CAAAGTCGTTGTCCGGCATTG	HR amplification primers for *TSN*; forward
3L:263300	GTGTTACTGATAAGGGCGTCAG	HR amplification primers for *TSN*; forward
TSN 5' Mut Primer #1	ACTCTGTGCGTGTGTGACcaacTGCACACGCTGTCCGATCACGCTGTCCG	Mutagenesis primer for *TSN* Binding Site 1
3L:264023	CTTCGAAAGTTCGTCCAAACTAAG	gRNA oligo cloning sequence *TSN* forward
3L:264045	AAACCTTAGTTTGGACGAACTTTC	gRNA oligo cloning sequence *TSN* reverse
Tbc1 primers		
Primer Name	Primer Sequence	Use of primer
CG4552_5L	GGGCGGCCGCCGACAAATAACACTTGTGGAAGTC	5' HR Region for *tbc1* ; forward
CG4552_5R	GGGGTACCTAAGGTTATTCAGCGGGATCTAAG	5' HR Region for *tbc1* ; reverse
CG4552_3L	GGGGCGCGCCTCATCACGTTCAAGTACGGATTC	3' HR Region for *tbc1* ; forward
CG4552_3R	GGCGTACGGTAGGTTCCCGTGTGCTTG	3' HR Region for *tbc1* ; reverse
CG4552 KO	gggttgggatgggttggggc	Diagnostics: Upstream of HR for *tbc1* KO; forward
DJ015	cacacacactcacaaaggagg	Diagnostics: Downstream of HR for *tbc1* KO; reverse
DJ066	cgcgcggcagccatatg ACAGAGTTTCCTGTGCCGGATG	Tbc1 cloning into pET15b protein expression; forward
DJ050	ccggatcctcgagcatatg CTACTTGGCGTCGTCAAGCACC	Tbc1 cloning into pET15b protein expression; reverse
DJ042	cacc ATGGAGGAGAATATGTGGATC	Tbc1 cloning into pENTRD vector; forward
DJ043	CTACTTGGCGTCGTCAAGCACC	Tbc1 cloning into pENTRD vector; untagged, reverse
DJ044	CTTGGCGTCGTCAAGCACCTGCA	Tbc1 cloning into pENTRD vector; forward

Continued on next page

White primers		
Primer Name	Primer Sequence	Use of primer
PW25Primer2	ACTGTGCGACAGAGTGAGAG	In w+ gene in pw25 construct; reverse
PW25Primer3	GGTCGACTCTAGAGGATCAT	In w+ gene in pw25 construct; forward
Sage primers		
Primer Name	Primer Sequence	Use of primer
Sage 5'	TCTCAACTACACCCTAACCGC	TOPO Cloning, UAS-sage-GFP (5' Primer)
Sage 3'	aacgatcattggggagcgtaa	TOPO Cloning, UAS-sage-GFP (3' Primer)
Sens primers		
Primer Name	Primer Sequence	Use of primer
CS00	cacc ATGAATCACCTATCGCCG	TOPO Cloning, UAS-sens-GFP (5' Primer)
CS01	GCAGCTGCTGCTGCTCACCTCCATC	TOPO Cloning, UAS-sens-GFP (3' Primer)
Fkh primers		
Primer Name	Primer Sequence	Use of primer
DJ060	GTCCTGGACTCGGCGGCGGCGGTCGcCAGCATGAGCGCCAGCATG	Forward; Site directed mutagenesis of fkh (t->c; underlined)
DJ061	TTCAGAATAACCCCACCAGAATGTTGaCGCTCTCGGAGATCTATC	Forward; Site directed mutagenesis of fkh (g->a; underlined)
DJ064	GTAGTCGGTGATCAGCGACgGATGGCACCGATTGGCCATCATC	Reverse; Site directed mutagenesis of fkh (adding in underlined)

Supplemental Table 2: List of Primers Designed. Sequence from target vector – small letters sequence change for mutagenesis – small letters, underlined

PAGE INTENTIONALLY TO BE BLANK

APPENDIX B

Antibody Supplemental Tables

.

Primary Antibodies

Antibody	Species	Use	Concentration	Source	RRID
Crb (CQ4)	Mouse	IHC	1:40	DSHB	AB_528181
CrebA	Rabbit	Polytene Chromosomes	1:5,000	Fox et al., 2010	AB_1085295
CrebA	Rabbit	IHC	1:5,000	Fox et al., 2010	AB_1085295
CrebA	Rat	Polytene Chromosomes		Andrew et al., 1997	
DCSP1 (ab49)	Mouse	IHC	1:100	DSHB	AB_2307345
en	Rabbit	IHC	1:100	Myat and Andrew, 2002	
GFP	Mouse	IHC	1:5,000	Molecular Probes	
GFP	Mouse	Western	1:5,000	Molecular Probes	
GFP	Rabbit	IHC	1:2,000	Life Technologies	A11122
GFP	Rabbit	IHC	1:10,000	Life Technologies	A11122
GM130	Rabbit	IHC	1:100	Abcam	
Golgin245	Goat	IHC	1:400	DSHB; Riedel et al 2016	
Rab11	Rabbit	Western	1:5,000	Yim and Andrew (unpublished)	
Rab11	Rabbit	IHC	1:500	Yim and Andrew (unpublished)	
Rab5	Guinea Pig	Western	1:10,000	Tanaka and Nakamura, 2008	
Sage	Rat	IHC	1:1,000	Fox et al, 2013	
SG2	Rabbit	IHC	1:8,000	Abrams et al., 2006	
Tbc1	Rat	Western	1:1,000	This work	
Tbc1	Rat	IHC	1:10	This work	
β-Gal	Mouse	IHC	1:100,000	Promega	
β-Gal	Rabbit	IHC	1:10,000	Promega	
βtub (AA4.3)	Mouse	Western/LICOR	1:1,000	DSHB	AB_579793
βtub (E7)	Mouse	Western/LICOR	1:500	DSHB	AB_2315513
βtub (E7)	Mouse	Western	1:1,000	DSHB	AB_2315513

Supplemental Table 2: List of Primary Antibodies.

Secondary Antibodies

Antibody	Species	Conjugate	Concentration	Source
Goat	Donkey	488	1:500	Invitrogen Molecular Probes
Guinea Pig	Goat	647	1:500	Invitrogen Molecular Probes
Guinea Pig	Goat	800CW	1:10,000	LI-COR
Mouse	Goat	488	1:500	Invitrogen Molecular Probes
Mouse	Goat	568	1:500	Invitrogen Molecular Probes
Mouse	Goat	Biotin	1:500	Invitrogen Molecular Probes
Mouse	Goat	HRP	1:10,000	
Rabbit	Donkey	555	1:500	Invitrogen Molecular Probes
Rabbit	Donkey	568	1:500	Invitrogen Molecular Probes
Rabbit	Goat	488	1:500	Invitrogen Molecular Probes
Rabbit	Goat	568	1:500	Invitrogen Molecular Probes
Rabbit	Goat	800CW	1:10,000	LI-COR
Rabbit	Goat	Biotin	1:500	Invitrogen Molecular Probes
Rat	Chicken	488	1:500	
Rat	Donkey	647	1:500	Invitrogen Molecular Probes
Rat	Goat	555	1:500	Invitrogen Molecular Probes
Rat	Goat	647	1:500	
Rat	Goat	Biotin	1:500	Invitrogen Molecular Probes
Rat	Goat	HRP	1:10,000	

Supplemental Table 3: List of Secondary Antibodies

References

Abrams, E.W., and Andrew, D.J. (2005). CrebA regulates secretory activity in the Drosophila salivary gland and epidermis. Development *132,* 2743-2758.

Andrew, D.J., Baig, A., Bhanot, P., Smolik, S.M., and Henderson, K.D. (1997). The Drosophila dCREB-A gene is required for dorsal/ventral patterning of the larval cuticle. Development *124,* 181-193.

Fox, R.M., Hanlon, C.D., and Andrew, D.J. (2010). The CrebA/Creb3-like transcription factors are major and direct regulators of secretory capacity. The Journal of Cell Biology *191,* 479-492.

Fox, R.M., Vaishnavi, A., Maruyama, R., and Andrew, D.J. (2013). Organ-specific gene expression: the bHLH protein Sage provides tissue specificity to Drosophila FoxA. Development *140,* 2160-2171.

Myat, M.M., and Andrew, D.J. (2002). Epithelial Tube Morphology Is Determined by the Polarized Growth and Delivery of Apical Membrane. Cell *111,* 879-891.

Riedel, F., Gillingham, A.K., Rosa-Ferreira, C., Galindo, A., and Munro, S. (2016). An antibody toolkit for the study of membrane traffic in Drosophila melanogaster. Biol. Open *5,* 987-992.

Tanaka, T., and Nakamura, A. (2008). The endocytic pathway acts downstream of Oskar in Drosophila germ plasm assembly. Development *135,* 1107-1117.

CURRICULUM VITAE FOR Ph.D. CANDIDATE
Johns Hopkins University School of Medicine

DOROTHY M. JOHNSON AUG. 1, 2018
Born: May 16, 1987, Bethesda, MD.

EDUCATION HISTORY

Ph.D. Expected	2018	Johns Hopkins School of Medicine, Baltimore, MD Biochemistry, Cell and Molecular Biology Mentor: Deborah J. Andrew, PhD
B.S.	2010	University of Maryland, College Park, MD Cell Biology and Genetics.
	2003	Montgomery College, Rockville, MD

OTHER PROFESSIONAL EXPERIENCE

Synthetic Biology Intern	2017	Revolve Biotechnologies, Inc., Baltimore, MD
Research Rotation	2011	Lab Erika Matunis, PhD, JHMI SOM
Research Rotation	2010	Lab of Geraldine Seydoux, PhD, JHMI SOM
Undergraduate Researcher	2008	Lab Eric Haag University of Maryland, College Park, MD.

HONORS and AWARDS

Ruth L. Kirschstein National Individual Research Award 5F31DE023721
Howard Hughes Medical Institute (HHMI) Undergraduate Fellowship

ACADEMIC HONORS AND AWARDS

2015	1st Place Speaker; BCMB Colloquium
Fall 2007	Semester Academic Awards, University of Maryland, College Park
Fall 2008	Semester Academic Awards, University of Maryland, College Park
Spring 2009	Semester Academic Awards, University of Maryland, College Park
Fall 2009	Semester Academic Awards, University of Maryland, College Park
Spring 2010	Semester Academic Awards, University of Maryland, College Park
2006	Honors Scholar Award, Montgomery College
Fall 2005	Dean's List, Montgomery College
Spring 2006	Dean's List, Montgomery College
Fall 2006	Dean's List, Montgomery College
Spring 2007	Dean's List, Montgomery College
2006	Phi Theta Kappa, honors society
2005	Millennium Scholar, honors society

PUBLICATIONS

Johnson, DM, Wells, MB, Lee, JS, Bastien, A, Slattery, M, and Andrew, DJ. CrebA directly regulates both regulators and effectors of secretory capacity. *In preparation*.

Johnson, DM, and Andrew, DJ. Drosophila Tbc1 is found in the trans portion of the *trans* Golgi Network and co-localizes to a subset of Rabs *In preparation*.

Beadell, A.V., Liu, Q., **Johnson, D.M.**, and Haag, E.S. (2011) Independent recruitments of a translational regulator in the evolution of self-fertile nematodes. *Proc. Natl. Acad. Sci USA* 108:19672-77. PMCID: PMC4316625

POSTERS AND ABSTRACTS

Johnson, DM, Andrew, DJ. (2012) The Role of tbc-1 in Drosophila Salivary Gland Development. Society of Developmental Biology 71st Annual Meeting, Montreal, Quebec, Canada, July 20 2012

Johnson, DM, Andrew, DJ. (2012) The Role of tbc-1 in Drosophila Salivary Gland Development. 54th Annual Drosophila Research Conference, Washington, DC, April 4, 2013

Johnson, DM, Andrew, DJ. (2014) The Role of tbc-1 in Drosophila Salivary Gland Development. 55th Annual Drosophila Research Conference, San Deago, CA, March 27 2014

Johnson, DM, Lee, JS, Wells, MB, Fox, R, Slattery, M, and Andrew, DJ. (2017) Salivary gland specific DNA binding of CrebA. 58th Annual Drosophila Research Conference, San Diego, CA, March 29 2017

TEACHING EXPERIENCE

Fall of 2012, 2013, and 2016

Teaching Assistant. Principles of Genetics. JHMI SOM